Elucidating the Role of MMS22L-TONSL Heterodimer in Homologous Recombination

DISSERTATION
ZUR
ERLANGUNG DER NATURWISSENSCHAFTLICHEN DOKTORWÜRDE
(Dr. sc. nat.)
VORGELEGT DER
MATHEMATISCH-NATURWISSENSCHAFTLICHEN FAKULTÄT
DER
UNIVERSITÄT ZÜRICH
VON

LUCIE JARMILA MLEJNKOVÁ

AUS DER
TSCHECHISCHEN REPUBLIK

PROMOTIONSKOMITEE:

PROF. DR. PETR CEJKA (VORSITZ UND LEITUNG DER DISSERTATION)

PROF. DR. MASSIMO LOPES

PROF. DR. MATTHIAS PETER

ZÜRICH, 2016

TO MY FAMILY

TABLE OF CONTENTS

ZUSAMMENFASSUNG

Die homologe Rekombination (HR) ist ein meist fehlerfreier DNS Reparaturmechanismus, welcher während der S-Phase des Zellzyklus stattfindet. HR wird verwendet um DNS-Doppelstrangbrüche (DSBs), DNS-Einzelstrangbrüche und sogenannte DNA-Vernetzungen zu reparieren. Der Mechanismus für die Reparatur von zwei-endigen DSB ist relativ genau bekannt, während weniger darüber bekannt ist, wie die Rekombination die DNS-Integrität nach Schäden, die durch die Verlangsamung und oder den Kollaps des Replisomes entstehen, aufrechterhält.

Das Hefe-Protein Mms22 ist speziell für die homologe Rekombination nach dem Abriss der DNS-Replikationsgabel verantwortlich. Die Homologe MMS22L (Mms22-like) und TONSL (Tonsoku-like) bilden einen Komplex in menschlichen Zellen. Nach Laser-induzierten DNS-Schäden lokalisiert der MMS22L-TONSL Komplex an ebendiese. Zellen, welche kein MMS22L oder TONSL besitzen, sind hypersensitiv für DNS-Schäden-induzierende Stoffe welche mit der DNS Replikation interferieren. Des Weiteren zeigen Zellen, welche kein MMS22L oder TONSL exprimieren, langanhaltende RPA Foki und weniger RAD51 Foki. Diese Erkenntnisse deuten darauf hin, dass der MMS22L-TONSL Komplex wichtig für die Initiation der HR-abhängigen Reparatur ist und eine Funktion nach der DNA-Resektion und vor dem Laden von RAD51 hat.

Das Ziel dieses Projektes war es, die mögliche Funktion des MMS22L-TONSL Komplex in der homologen Rekombination, vor allem als Rekombination-Mediators mittels biochemischen Analysen zu untersuchen. Rekombination-Mediatoren wie zum Beispiel BRCA2 oder RAD51 Paraloge spielen eine Rolle beim Laden von RAD51 auf RPA-bedeckte einzelsträngige DNS. RAD51 sucht danach homologe DNS Abschnitte was im Weiteren zu HR Reparatur führt.

Als erstes haben wir den MMS22L-TONSL Komplex in *Sf9* Insektenzellen exprimiert. Der Komplex liegt als Heterodimer vor. Wir untersuchten die Fähigkeit des Komplexes an DNS-Substrate zu binden und fanden, dass MMS22L-TONSL an lange DNS Moleküle bindet.

MMS22L-TONSL bindet auch an einzelsträngige DNS welche schon von RPA gebunden ist.

Wir testeten direkte Protein-Protein Interaktionen zwischen MMS22L-TONSL oder MMS22L und RAD51. Erstaunlicherweise hat RAD51 sowohl MMS22L-TONSL als auch MMS22L gebunden, was darauf hinweist, dass MMS22L die interagierende Untereinheit ist. MMS22L-TONSL erleichtert den RAD51-abhängigen DNS Austausch, welcher ATP benötigt. Dies geschieht durch Verminderung der Bindungsaffinität von RAD51 zu doppelsträngiger DNS. Folglich, bindet RAD51 eher an einzelsträngige DNS als an doppelsträngige DNS, wenn beide Substrate vorliegen. Unsere Ergebnisse deuten daraufhin, dass MMS22L-TONSL als Rekombination-Mediator in der homologen Rekombination wirkt, ähnlich wie BRCA2 und RAD51 Paraloge wie RAD51B, RAD51C, RAD51D, XRCC2 und XRCC3.

SUMMARY

Homologous recombination (HR) is a mostly accurate DNA repair pathway, which takes place in the S-phase of the cell cycle. It repairs DNA double-strand breaks (DSBs), DNA single-strand breaks and contributes to the repair of interstrand DNA crosslinks. While the mechanism of a two-ended DSB is relatively well understood, how recombination restores DNA integrity upon replisome stalling or collapse, remains poorly understood.

In yeast, Mms22 was found to promote homologous recombination specifically upon DNA replication fork stalling. In human cells, the human homologue MMS22L (Mms22-like) forms a complex with TONSL (Tonsoku-like). MMS22L-TONSL complex localizes at sites of damaged DNA upon laser microirradiation. Depletion of MMS22L or TONSL caused hypersensitivity to a variety of DNA-damaging drugs that interfere with DNA replication. Furthermore, depletion of MMS22L or TONSL resulted in persistent RPA foci and decreased formation of RAD51 foci. These data suggested that MMS22L-TONSL complex might play a role in the initiation of HR-dependent repair specifically downstream of DNA end resection prior to RAD51 loading.

The aim of this project was to perform biochemical analyses to investigate the potential role of MMS22L-TONSL in HR, in particular as a recombination mediator. Recombination mediators, such as BRCA2 or RAD51 paralogs, help to load RAD51 on RPA-coated ssDNA. RAD51 then performs search for homologous DNA duplex, which subsequently leads to HR repair.

First, we purified MMS22L-TONSL complex from *Sf*9 insect cell expression system. The complex purified as heterodimer. We then tested its ability to bind various DNA substrates. We found that MMS22L-TONSL complex binds long stretches of DNA, suggesting that DNA might be bound cooperatively. Furthermore, MMS22L-TONSL complex bound RPA-coated ssDNA. Next we tested direct protein-protein interactions between MMS22L-TONSL or MMS22L and RAD51. Strikingly, RAD51 binds both MMS22L-TONSL and MMS22L, suggesting that MMS22L is the interacting subunit. MMS22L-TONSL also facilitates RAD51-dependent DNA strand exchange activity in an ATP-dependent manner by lowering RAD51's binding affinity to dsDNA. As a result, this facilitates RAD51 filament formation on single-stranded DNA when dsDNA is present. Our data collectively support a notion that MMS22L-

TONSL acts in homologous recombination as a recombination mediator, similarly as BRCA2 and RAD51 paralogs such as RAD51B, RAD51C, RAD51D, XRCC2, and XRCC3.

1 INTRODUCTION

1.1 HOMOLOGOUS RECOMBINATION

Homologous recombination (HR) is a high fidelity repair pathway that recognizes and repairs DNA lesions including double-strand DNA breaks (DSBs), single-strand DNA (ssDNA) breaks, and contributes to the repair of interstrand DNA crosslinks in S and G2 phases of the cell cycle.[1,2] In contrast, the non-homologous end-joining (NHEJ) pathway is responsible for repairing DSBs throughout the cell cycle by ligating the processed DNA ends. HR copies the missing DNA information from homologous DNA regions, such as sister chromatids, resulting in a highly accurate repair mechanism.[2]

Homologous recombination is initiated by 5' DNA end resection; this step requires three nucleases, including MRE11-RAD50-NBS1 (MRN complex), Exonuclease I (Exo1), and DNA2 in conjunction with helicases Bloom syndrome protein (BLM) or Werner syndrome helicase.[3,4] First, MRN complex binds to the DSB and serves as a sensor of a DSB and signals it to the downstream factors as well. Then, the DSBs are processed by MRN together with CtIP [5,6] creating a 3' ssDNA overhang to be further processed by either BLM-DNA2 or WRN-DNA2 together or with Exo1 and resulting in a long-range DNA end resection.[3,6] Those extensively processed 3'-ends are rapidly covered by the replication protein A (RPA) to prevent formation of secondary structures and protect the naked DNA.[7] RPA is then replaced with recombinase RAD51. Since RPA has an inhibitory effect on RAD51 binding to ssDNA, this step requires recombination mediators, which are the focus of this thesis. RAD51-coated ssDNA filament (presynaptic filament) is later stabilized by RAD54. RAD51 then performs strand invasion into the sister chromatid and searches for homologous region in a rather random manner by testing various segments of the donor DNA duplex. Once the 3'-end ssDNA overhang is paired to an appropriate region, it forms a displacement loop (D-loop). RAD54 then facilitates RAD51 dissociation from ssDNA, allowing DNA polymerases to fill up the gap. There are two sub pathways in DSB repair, synthesis-dependent strand annealing (SDSA) and double-strand break repair pathway (DSBR). They both share the initial steps of HR but differ in downstream steps of RAD51 filament formation. In SDSA, once the gap is synthetized, the invading

strand is re-annealed back to the original DNA strand before the breakage. This sub pathway produces only non-crossover products. In DSBR, on the other hand, Rad52 captures the second end of the broken DNA and anneals it to the D-loop structure, forming a double Holliday junction (dHJ). Once the gaps are filled, dHJ structure is later resolved into crossover (CO) or noncrossover (NCO) products (Figure 1).[1,2,8,9,4]

There are three pathways, which evolved to process dHJs. First, dissolution pathway involves BLM-TOPOIIIα-RMI1-RMI2 (BTR) complex, which is employed in the S phase of the cycle and generates exclusively NCO products to prevent loss of heterozygosity in mitotic cells. The second, resolution pathway, is catalyzed by the SLX1-SLX4-MUS81-EME1 (SLX-MUS) complex and results in the mixture of both CO and NCO products. The third involves GEN1, an enzyme functioning in resolution pathway to process dHJs in case of missing function of BTR and/or SLX-MUS enzymes.[10,11]

In meiotic cells, DSBs are introduced intentionally in order to generate crossover products subsequently resulting in genetic diversity. Such events require factors specific for meiotic recombination, such as Dmc1 (disrupted meiosis cDNA 1) and Hop2-Mnd1 complex, which will be further discussed below in the text. Dmc1 is only expressed in meiosis and shares up to 50% sequence similarity with Rad51; in fact, both Rad51 and Dmc1 are homologs of bacterial RecA, mentioned below. Both Rad51 and Dmc1 bind to 3' resected ssDNA ends of the DSB and with help of Hop2-Mnd1 heterodimer generate the D-loop, the first recombination intermediate.[12-14]

On the contrary to eukaryotes, in bacteria *Escherichia coli* (*E. coli*), the DNA breaks are mainly repaired by homologous recombination. Moreover, two overlapping pathways evolved, namely the RecBCD and RecF pathways. RecBCD is employed when dsDNA breaks occur, whereas the second pathway, RecF, primarily repairs single-stranded DNA gaps that arise during replication process of a damaged DNA template. RecF can also repair dsDNA breaks as a compensatory pathway to RecBCD. [15] In the RecBCD pathway, once the DSB is detected, the RecBCD helicase/nuclease binds to the DSB and starts processing the ends. The RecB motor translocates in a 5' to 3' direction on the ssDNA, whereas RecD motor is sliding in 3' to 5' direction on the opposite strand but in the same overall direction.[16,17] RecBCD then loads recombination protein A (RecA; RAD51 homolog) onto 3'-tailed ssDNA and then dissociates from the RecA-coated ssDNA filament, showing that resection and

recombinase-loading steps are coupled. In the RecF pathway, the repair of an ssDNA gap starts with a 5' to 3' resection by RecJ nuclease (Exo1 homolog) in conjuction with a 3' to 5' RecQ helicase. The ssDNA regions are covered with single-strand DNA binding protein (SSB; RPA homolog). The RecFOR complex consists of 3 subunits, RecF, RecO and RecR and mediates the nucleation of RecA onto SSB-coated ssDNA filament while displacing SSB at the same time. Both RecBCD and RecF pathways employ RuvA, RuvB and RuvC to resolve Holliday junction.[15,18,19]

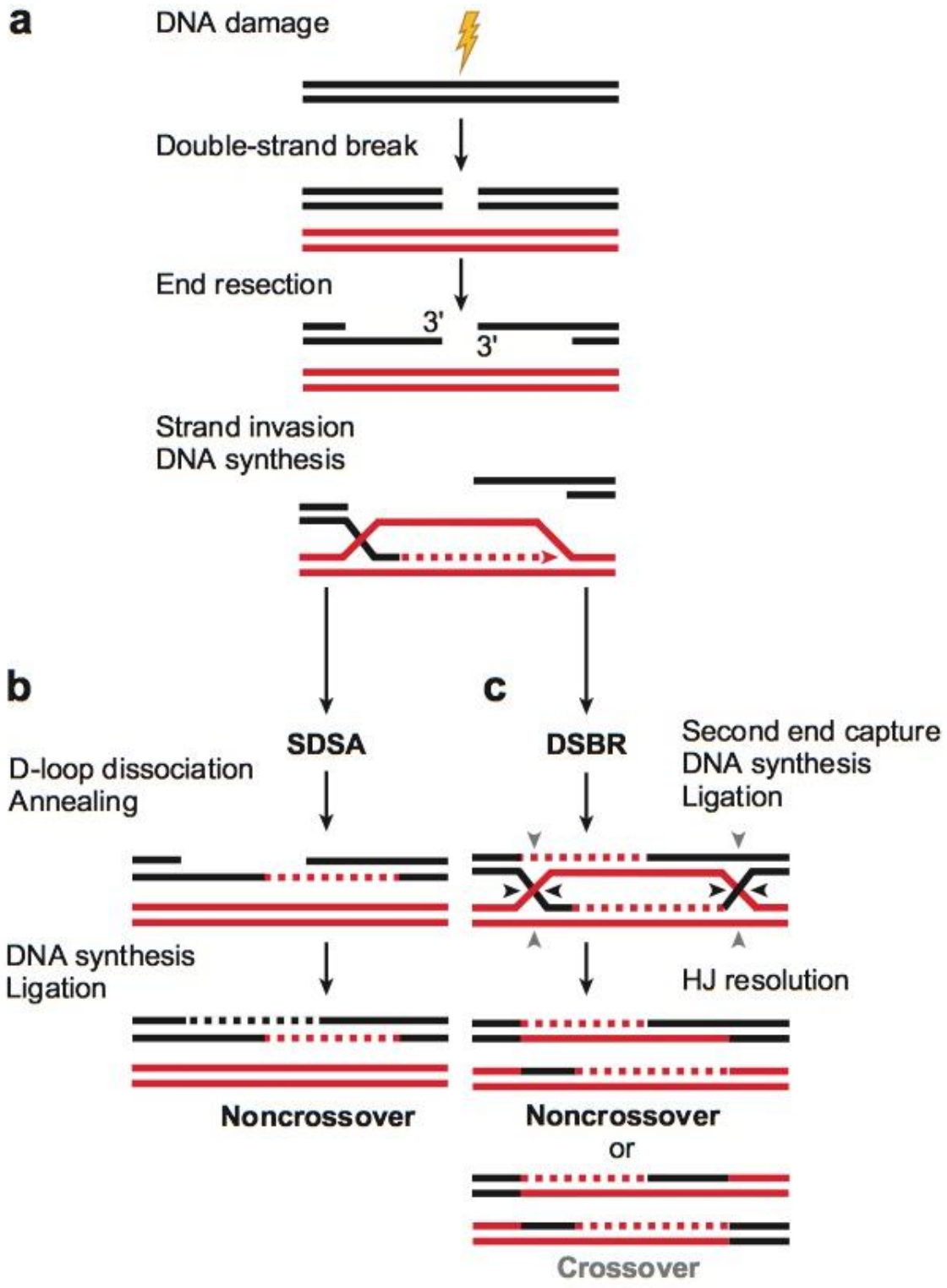


Figure 1: Pathways of DSB repair in homologous recombination. [2]

1.2 REGULATION OF HOMOLOGOUS RECOMBINATION

1.2.1 Rad52

Rad52 is a key protein in HR in *Saccharomyces cerevisiae*. In budding yeast, Rad52 loads RAD51 onto RPA-coated ssDNA filament[20,21] and anneals complementary ssDNA coated with RPA. The first of yeast Rad52 roles is similar to that of BRCA2 in mammals. [20,21] The second role is conserved in evolution and lies in its necessity in the annealing step where the second end is captured in the single-strand annealing pathway. [22] Knowing the importance of Rad52 in budding yeast, it was surprising that *RAD52* knockout mice were not affected in homologous recombination [23] as well as in *U. maydis* and chicken cells. [24] However, its importance in HR became evident in *BRCA2*-deficient human cell line where human RAD52 inactivation leads to defective HR and decrease of RAD51 foci formation upon treatment with ionizing radiation (IR). [25] Human RAD52 lacks recombination mediator activity in *in vitro* biochemical assays, therefore it is thought that it possibly synergizes with RAD51 paralogs (RAD51B, RAD51C, RAD51D, XRCC2, and XRCC3), which will be described below. A hint leading to such assumption is provided by a study where DT40 chicken cells were defective in XRCC3 function after *RAD52* inhibition.[21,24]

1.2.2 BRCA2

RAD51 foci accumulate at ssDNA regions formed after resection of broken DNA ends. However, to localize at those sites, RAD51 focus formation requires specific HR factors, known as recombination mediators, including RAD51B, RAD51C, RAD51D, XRCC2, XRCC3, and BRCA2 [32], which will be discussed below. Recombination mediators have indispensables roles in the HR, where they promote the formation and stabilize RAD51 filaments on ssDNA through multiple mechanisms.

A fraction of breast or ovarian cancers is associated with hereditary or sporadic mutations in *BRCA1* [26], *BRCA2* (breast cancer type 2 susceptibility protein) or *PALB2* (partner and localizer of BRCA2) genes. Mutations in these genes lead to excessive genome instability.[27,28] Moreover, Brca1-deficient mice exhibit early embryonic lethality, similarly to Brca2.[29] *In vivo*, BRCA1 and BRCA2 were shown to interact with

each other, however the link between them was unclear. The study conducted by Sy et al. [27] describes PALB2 as a scaffolding protein amongst BRCA1 and BRCA2. PALB2 was shown to directly bind to BRCA1 via the coiled-coil motif and also serves as a bridging protein to the BRCA2-RAD51 complex. PALB2 interacts with BRCA2 via its C-terminal domain, specifically with its transactivation core sequence. Furthermore, all three proteins, BRCA1, BRCA2, and PALB2 colocalize at sites of DNA damage as nuclear foci upon DNA damage. Interestingly, BRCA1 acts upstream of PALB2 as BRCA1 foci formation is unaffected by disrupting the interaction between these two proteins. Conversely, PALB2 is a compulsory factor for the recruitment of BRCA2 to DNA damaged sites. [28,30] In accord, BRCA2-defective CAPAN-1 cell line shows inability of RAD51 to form nuclear foci, however such phenotype was not observed for BRCA1, where the foci at DNA damage sites formed normally.[31]

The breast cancer tumor suppressor BRCA2 is a key recombination mediator in homologous recombination in high eukaryotes. Mutations in the *BRCA2* gene predispose affected individuals to high risk of ovarian and breast cancer [27,28], pancreatic cancer, and other tumors in infants suffering of Fanconi anemia. [33]

While in budding yeast BRCA2's function is supplemented by Rad52 and BRCA2 homolog is missing completely as mentioned above, several BRCA2 homologs can be found in other species, such as BRCA2 in mice (*Mus musculus)*, BRC-2 in the nematode *Caenorhabditis elegans*, or Brh2 in the fungus *Ustilago maydis.*[34] Human BRCA2 is structured into several domains. The N-terminal region is interacting with PALB2 at amino acids between 21 and 39. The BRC1-8 repeats at amino acid region between 1009 and 2083 are RAD51 binding sites together with the C-terminal TR2 region.[35,36,37] BRCA2 also possesses DNA-binding domain consisting of a helical domain (H), three oligonucleotide binding (OB) folds and a tower domain (T), possibly necessary for binding to both ssDNA and dsDNA.[35,38] Human BRCA2 also has a Dmc1 interaction site, known as PhePP motif, a 26-amino-acid residue region located between amino acid residues 2386 and 2411 (Figure 2).[2] BRCA2 interacts with human RPA through its N-terminal region located at the amino acid region between 18-105, called an acidic transactivation domain.[39] Furthermore, a cancer predisposing mutation Y42C in *BRCA2* affected RPA binding, pointing at the importance of interaction.[40]

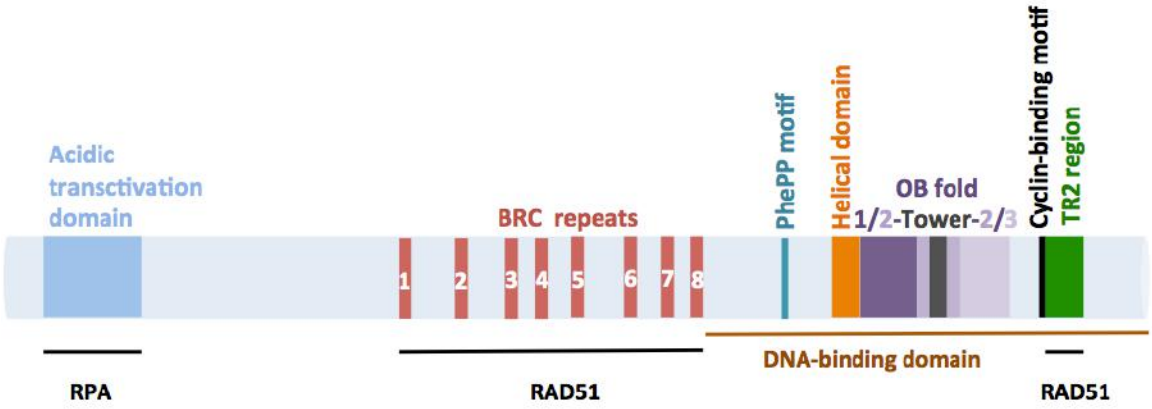


Figure 2: BRCA2 domain structure.

BRCA2 orchestrates the HR machinery through the interaction with the RAD51 recombinase. It helps to overcome inhibitory effect of RPA-bound DNA and promotes the nucleation and growth of RAD51 on the ssDNA overhang. RAD51 is known to hydrolyze ATP upon DNA binding. The DNA-dependent ATPase activity of RAD51 is regulated by the Mg^{2+} and Ca^{2+} cations. *In vitro*, the presence of Mg^{2+} results in a fast ATP hydrolysis, meaning that ATP (RAD51-ATP-ssDNA; active form) is quickly converted into ADP (RAD51-ADP-ssDNA; inactive form). On the other hand, addition of Ca^{2+} slows down the ATP hydrolysis rate, therefore keeping the active form of the filament, i.e. RAD51-ATP-ssDNA. Interestingly, such an effect has not been shown in the yeast Rad51 counterpart. [41]

Initial studies focused on BRCA2 mechanistic function in HR and repair of dsDNA breaks were hampered by the difficulties in purifying the full-length protein due to its size, 3418 amino acids (~384 kDa).[34,42] In 2010, however, three different groups purified independently the full-length protein and provided insights into BRCA2 function *in vitro*.[20,34,42] Purified full-length BRCA2 from human cells binds both ssDNA and dsDNA oligonucleotide substrates [20], but shows a strong preference for ssDNA substrates, including 3' tailed and 5'tailed DNA oligonucleotides. BRCA2 functions in the HR repair as a recruitment factor of RAD51, which brings RAD51 to the resected ssDNA regions. This is a crucial reaction step in the HR as impairment of BRCA2-RAD51 interaction leads to chromosome instability and a failure to repair dsDNA breaks in S-phase.[43]

RAD51 catalyzes the strand exchange reaction and searches for homology regions within the homologous DNA of the sister chromatid. Under experimental conditions, where RPA is not covering the 3' ssDNA overhang, RAD51 can perform the strand exchange reaction by itself. Additionally, RAD51 must be pre-bound to ssDNA first and then dsDNA is added to the reaction.[20] RAD51 was shown to have an affinity to both ssDNA and dsDNA, which is detrimental for strand exchange reaction as binding of RAD51 to dsDNA is inhibitory and strand invasion cannot occur. BRCA2 possesses the ability to tackle both situations - it inhibits RAD51 binding to dsDNA and directs it to ssDNA, where it helps its loading onto RPA-coated ssDNA. [20,34,42]

What is the exact mechanism of BRCA2 mediator function in promoting RAD51 nucleation on ssDNA? BRCA2 possesses eight BRC repeats, BRC1-8. Crystallographic studies revealed that BRC4 repeat binds RAD51 through several conserved interaction sites, namely Phe 1524, Ala 1527, Leu 1545, and Phe 1546.[45] This was later shown in *in vitro* assays where BRC1-8 and BRC4 interactions with RAD51 and DNA substrates were investigated.[46]

The binding of RAD51 to dsDNA can be reduced by the BRC4 repeat alone. Similarly, BRC4-bound BRCA2 stabilized RAD51-ssDNA nucleofilament by preventing RAD51 ATP hydrolysis and keeping RAD51 in its ATP-bound form. This is necessary for strand invasion in later stage of HR.[46] Two modes of action were identified for the two distinct BRC repeat regions. BRC1-4 binds free RAD51 unlike BRC5-8 that does not have such ability. BRC5-8, on the other hand, binds RAD51-coated ssDNA with a high affinity.[37,47] This is in agreement with previous studies where full-length BRCA2 was reported to bind 4-6 RAD51 molecules.[20,42]

In Chatterjee et al. [37] study, the BRC5-8 region (1596 to 2152 amino acid region) coupled to DNA-binding domain located at 2153 to 3418 amino acid region) (DBD-BRC5-8) revealed a distinct function of BRC5-8 to that of BRC1-4 in promoting RAD51 strand exchange activity. And it was also shown that DBD-BRC5-8 binds, thus stabilizes, RAD51-coated filament with ~1.5-fold higher affinity than DBD-BRC1-4 and DBD-BRC1-8 region.[37] In summary, BRC1-4 promotes nucleation of RAD51 onto ssDNA by reducing RAD51's ATPase activity as discussed above, whereas BRC5-8 repeats stabilize the growing RAD51-ssDNA filament and promote RAD51 strand exchange activity.

Moreover, there is evidence that BRCA2 functions not only in HR to repair DSBs, but it is also involved in the protection of stalled DNA replication forks by averting the degradation of newly synthesized DNA strand by MRE11. As mentioned above, BRCA2 possesses nine interacting sites with RAD51; eight BRC repeats in the central part and TR2 domain (also known as C-ter) located at the C-terminus. While BRC repeats are necessary for DSB repair, it has been shown that TR2/C-ter is required for replication fork protection. Specifically, if the BRCA2 is missing the TR2/C-ter, MRE11 degrades the nascent strand of a stalled DNA replication fork. Such degradation can be reduced by introduction of the MRE11-inhibitor mirin. Therefore BRCA2's C-terminal TR2/C-ter is involved in the protection of stalled replication forks, a mechanism differing from the traditionally known DSB repair[33], and the function of which has yet to be determined.

1.2.3 RAD51 PARALOGS

In *Saccharomyces cerevisiae*, Rad55 and Rad57 were initially the only two RAD51 paralogs identified in the budding yeast. However, later it was shown that two other proteins, Csm2 and Psy3, belong to the Rad51 paralogs family.[21,48,49] Rad55 and Rad57 form a heterodimer, while Csm2 and Psy3 bind two Shu proteins (Shu1 and Shu2), resulting in the Shu complex.[49] The Rad51 paralogs from the budding yeast are, similarly as human Rad51 paralogs, essential in recombination. Mutations in these proteins lead to chromosomal instability.[50] Rad55-Rad57 is a mediator protein that forms a complex with Rad51 bound to ssDNA and stabilizes Rad51-ssDNA filament by inhibiting the Srs2 translocase/helicase activity, which would be otherwise stimulated by Rad51 coated ssDNA filament formation. Therefore Rad55-Rad57 protects newly nucleated Rad51 filament on the ssDNA.[21] However, to allow Rad51 to nucleate onto RPA-coated ssDNA, Rad51 requires help of additional factors to overcome the inhibitory effect of RPA. *In vitro*, the Shu complex together with Rad55-Rad57 and Rad52 promote Rad51 strand exchange reaction in the presence of RPA. Mutation in the *csm2-F46* allele impairs interaction of the complex with

Rad55 resulting in an inability to promote Rad51-ssDNA filament formation *in vitro* and in a defective HR in *vivo.*[49]

Alongside BRCA2, five other human recombination mediators, all members of the *RAD51* gene family, are involved in HR. These are *RAD51B* (*RAD51L1*), *RAD51C* (*RAD51L2*), *RAD51D* (*RAD51L3*), *XRCC2*, and *XRCC3.*[50] Individual mutants of each RAD51 paralog are severely impaired in RAD51 focus formation upon DNA damage induced by ionizing radiation (IR), pointing towards their crucial role in HR.[51] RAD51 paralogs share around 20-30 % identity of amino acids with RAD51, however all are lacking recombinase activity *per se.*[52,53] Mutants generated in chicken DT40 cells, *rad51c*, *rad51d*, *xrcc2*, and *xrcc3* were all defective in RAD51 focus formation and were sensitive to cisplatin treatment. This was partially overcome by overexpressing RAD51, which could to some extent supplement the missing paralog's function.[51] This suggested that the RAD51 paralogs function in conjunction with RAD51. Interestingly, individual *rad52* and *xrcc3* mutants were viable and displayed no hindrance in the growth, however the double mutant *rad52 xrcc3* was non-viable.[24] None of the paralogs self-interacts such as RAD51, which forms heptamers in solution.[2,52] However, the paralogs form several different complexes within each other. In HeLa cells, RAD51B-RAD51C-RAD51D-XRCC2 (BCDX2) and RAD51C-XRCC3 (CX3) were detected. Human recombinant RAD51C and XRCC3 interact directly as well as RAD51B and RAD51C.[54,55] Furthermore, RAD51D forms a complex with XRCC2.[56] Shu1 is a homolog of XRCC2 and Psy3 is a homolog of RAD51D. Furthermore, Shu2 is homologous with human SWS1, a protein, which interacts with RAD51, as well as with RAD51 paralogs.[57] The function of SWS1 in human cells is yet to be determined.

Depletion of XRCC3 results in reduced HR, but RAD51 foci form normally at sites of DNA damage. RAD51 focus formation was also unaffected in XRCC3$^{-/-}$ colorectal cells. Furthermore, simultaneous depletion of RAD51C and XRCC3 (CX3) did not affect RAD51 focus formation either, suggesting a role for CX3 complex downstream of RAD51 binding to ssDNA. In accord, after induction of DNA damage, XRCC3 foci colocalized with RAD51, however they persisted even after the disappearance of RAD51 foci.[57]

RAD51B and RAD51C form a heterocomplex and represent the known human homologs of the Rad55-Rad57 recombination mediator complex found in the budding yeast *S. cerevisiae.*[55] Coprecipitation experiments revealed that RAD51B and RAD51C form a stable protein complex and recombinant human RAD51B and RAD51C proteins purified from *Sf9* insect cell interact directly as well.[54] Similarly to BRC1-4, purified RAD51B-RAD51C complex binds ssDNA with a high affinity and shows low binding to dsDNA. Furthermore, RAD51B-RAD51C shows ATPase activity upon binding to ssDNA and promotes RAD51 strand exchange activity, however the complex alone cannot overcome RPA inhibitory effect sufficiently. This step likely requires other RAD51 paralogs.[54] Purified BCDX2 complex was shown to preferentially bind ssDNA, specifically ssDNA nicks in dsDNA and displays ATP-hydrolysis activity as RAD51. Furthermore, RAD51C-RAD51D complex interacts with XRCC2 via RAD51D in yeast two-hybrid system.[58,59] RAD51D also binds DNA and possesses ATPase activity stimulated upon binding to ssDNA and dsDNA and forms a heterodimer with XRCC2.[53,59] Interestingly, in *Caenorhabditis elegans*, Rad51 paralog complex RFS-1/RIP-1 binds ssDNA but not dsDNA and also stimulates Rad51 strand exchange *in vitro*. *In vivo* analyses showed a necessity of RFS-1/RIP-a for Rad51 foci formation in HR [60], which is reminiscent of human RAD51 paralogs.

1.2.4 RAD54

RAD54 translocase is a protein belonging to the *RAD52* epistasis group, a class of enzymes sharing common features, such as mutation in any of them results in defects in HR, as well as in sensitivity to DNA-damaging agents.[8,61] RAD54's function in HR has mainly been characterized biochemically. RAD54 is a multifunctional translocase involved in all three steps of HR, pre-synapsis, synapsis, and post-synapsis.[8] In pre-synapsis, RAD54 displaces RAD51 from dsDNA in order to allow Rad51 to bind ssDNA instead. In synapsis, RAD54 stimulates DNA strand exchange in the presence of Ca^{2+}, stimulates DNA heteroduplex extension promoted by RAD51, and also can translocate along dsDNA using the energy of ATP hydrolysis. And finally, in post-synapsis, Rad54 promotes branch migration of Holliday junction in an ATP-dependent manner.[62] *In vivo*, RAD54 forms nuclear foci at sites of DNA damage

and colocalizes with RAD51 at such sites.[63] Mouse embryonic stem cells lacking *RAD54* are sensitive to ionizing radiation and display defects in HR, similarly to yeast *RAD54* mutants. On the other hand, *RAD54* knockout mice are not sensitive to IR treatment, but show sensitivity to mitomycin C, an interstrand DNA cross-linking agent.[63] In humans, there two RAD54 homologs, RAD54 and RAD54B.[14] The data suggest that their function in HR is likely redundant as they are both double-stranded DNA-dependent ATPases, can translocate along DNA, and display double-helix-opening activities.[63] RAD54B is thought to be a homolog of yeast Rdh54 (also known as Tid1) based on amino acid sequence homology. However, functional studies hint towards a possibility that this is not a true RAD54B homolog.[63] Rdh54/Tid1 interacts with Dmc1 during meiosis as shown for RAD54B.[14,64] Biochemically, Rdh54 possesses ATPase activity and similarly as RAD54 promotes RAD51 strand invasion/D-loop formation. However, *rad54* and *rdh54* phenotypes differ in response to DNA damaging agents. *rad54* mutants are sensitive to the treatment with alkylating agent methyl methansulfonate (MMS) unlike *rdh54*. Furthermore, *RAD54* and *RDH54* likely have to some extent overlapping function in meiosis, as double mutants *rad54 rdh54* resulted in lower sporulation efficiency and spore viability in comparison to individual mutants.[63]

1.2.5 Hop2-Mnd1 complex

Dmc1 and Rad51 are both necessary factors for successful meiotic recombination. Unlike Rad51, Dmc1 is only expressed in meiosis. Dmc1 is known to form an octameric complex, resembling a ring and such structure is then bound to Rad51-coated ssDNA filament. In fact, Rad51 is required in order to load Dmc1 to a resected ssDNA overhang as *rad51* mutants resulted in an inability of Dmc1 to bind to the sites of the breaks. Similarly to Rad51, Dmc1 can also catalyze strand exchange reactions *in vitro*, resulting in a D-loop formation.[14] Rad51/Dmc1 together with Hop2-Mnd1 promote strand invasion and align the 3' ended ssDNA overhang to a homologous region. Rad51/Dmc1 can catalyze such reaction alone, however in a rather slow and ineffective manner.[12] Hop2-Mnd1 complex is a known evolutionarily conserved cofactor of the Dmc1 recombinase in meiotic cells in *S. cerevisiae*. Furthermore, it has been shown that mutations in the *HOP2* and *MND1* genes result

in a similar phenotype as *DMC1*, i.e. inability to promote DNA pairing (synapsis) between homologs, and such phenotype can be partially rescued by overexpressing Rad51.[12,14] Recent studies, however, revealed Hop2-Mnd1 importance in somatic cells in higher eukaryotes as well.[65] Hop2 and Mnd1 form a Y-shaped heterodimer possessing two dsDNA-binding domains at the N-terminus and one ssDNA-binding domain at the C-terminus of the complex. Two RAD51/DMC1 interaction sites are located at the C-terminus as well on both protein complex subunits. Hop2-Mnd1 complex binds with its C-terminus to ssDNA and via RAD51/DMC1 interaction sites stabilizes the RAD51-ssDNA filament. The V-shaped N-terminus of the complex then captures the dsDNA and forms a synaptic complex enabling RAD51 to search for homology regions within captured DNA duplex.[65]

1.3 MMS22L-TONSL

1.3.1 YEAST Mms22

First reports about *Saccharomyces cerevisiae* Mms22 are dated back to 1977. In this study, Prakash and Prakash [66] performed screening that identified 22 mutants sensitive to methylmethane sulfonate (MMS), an alkylating agent known for blocking DNA replication forks. Whilst other *mms* mutants were also sensitive to UV, X-rays, or both, five mutants showed sensitivity only to MMS, namely *mms1-4*, *mms2-1*, *mms4-1*, *mms5-2*, and *mms22-1.* [66] However, such mutants were not of any interest in that study, therefore the *mms1* and *mms22* mutants escaped further attention for many subsequent years.

Mms22 together with cullin Rtt101 and mms1 form a Cul4(Dbd1)-like E3 ubiquitin ligase [67], where it serves as substrate-specific adaptor. Mms22 is linked to Rtt101 via Mms1, similarly as human DDB1 (damage-specific DNA binding protein 1).[67,68]

However, later, Duro et al. [69] uncovered that Mms22 might be acting as a repair factor in HR. Based on these studies it is not clear, whether Mms22 is a link between HR and histones. Below, I will focus mainly on Duro et al. [69] study, which is relevant to my research.

The sensitivity of *mms22* and *mms1* mutants to agents blocking replication fork progression suggested their potential involvement in the repair of stalled or broken replication forks. To this point, Duro et al. [69] investigated Mms22 and Mms1 mutants and their role within the HR repair pathway. HR needs a template to copy the missing information from. Such scenario happens only in the S-phase or the G2-phase of the cell cycle. Indeed, in cells synchronized in G1-phase and then released into S-phase, those cells lacking MMS1 and MMS22, as well as RAD51 showed marked reduction of HR (unequal sister chromatid exchange; uSCE) after treatment with MMS in comparison with parental cells. Similar effect was observed when those cells were treated with camptothecin (CPT), a topoisomerase I inhibitor that causes ssDNA breaks and stalls DNA replication forks.[70] DNA template for HR can be found not only in sister chromatids but also in homologous chromosomes. [69] Treatment with different concentrations of MMS resulted in a lower frequency of heteroallelic recombination in cells lacking *MMS22* and *MMS1*. Therefore MMS22 and MMS1 function upon replication fork stalling, and are required for HR repair utilizing as a template either sister chromatid or homologous chromosome.

MMS22 and *MMS1* are clearly involved in the repair of ssDNA breaks but are they also required for HR induced by dsDNA breaks formed upon DNA damage? To answer this question, Duro et al. [69] went on and treated the cells with galactose that induced HO endonuclease, which subsequently created DSBs. As expected, dramatic increase in the frequency of HR was observed in wild-type cells, which could repair the dsDNA break but in contrast very low frequency in cells lacking *RAD51*. But surprisingly, in cells lacking *MMS22* and *MMS1*, the frequency of HR was unaffected. These results were further confirmed by treating the same mutants with ionizing radiation (IR), which is known for giving rise to DSBs.[69,71] Indeed, *MMS1* and *MMS22* lacking cells were unaffected by this treatment, resulting in a similar phenotype as described above, i.e. frequency of HR was almost as high as in wild-type cells. Collectively, these results indicate a possible function of *MMS1* and *MMS22* upon replication fork stalling but not in the HR initiated by DSBs.[69] However, the exact mechanism of the function of Mms22 and Mms1 remains to be elucidated.

1.3.2 HUMAN MMS22L-TONSL

Homologous recombination (HR) represents an essential process that is required for the progression and completion of DNA replication of mammalian cells. Although HR has been mostly studied in the context of direct DSB repair, these lesions are rather rare and are primarily repaired by non-homologous end joining pathway.[1] In contrast, the primary role of HR is the repair of single-strand DNA breaks (SSBs) encountered by DNA replication machinery, to restart stalled replication forks and to repair ssDNA gaps behind replication forks. These lesions cannot be repaired by non-homologous end-joining mechanisms, and are thought to be much more frequent than direct DSBs.

As already mentioned in the previous chapters, mechanistically, HR is initiated by 5' to 3' resection resulting in large 3'-terminated ssDNA overhang, which is then covered with RPA to prevent the formation of secondary structures. However, RPA-coated ssDNA filament results in an inhibitory effect on RAD51, showing the necessity of a recombination mediator to help load RAD51 onto RPA-coated ssDNA filament. Below, I will discuss recent studies [69,71-73], which show that MMS22L-TONSL might represent a novel, yet uncharacterized recombination mediator that is distinct from factors described in chapter 1.2.

1.3.2.1 Identification of the MMS22L-TONSL complex

Given the fact that Mms22 in yeast is a crucial part of the repair machinery initiated upon replication fork perturbation/replication fork stalling, it was intriguing whether this protein is evolutionarily conserved. In 2010, four independent groups described the identification of a human homolog, MMS22L.[71,72,73,74] Specifically, Duro et al. [71] performed multiple sequence alignment analyses and found an evolutionarily conserved region between yeast Mms22 (781-920) and human (373-535) C6ORF167. Since this was the only similarity found between these two proteins, the human ortholog was called Mms22-like (MMS22L). Furthermore, search for potential interacting partners of MMS22L by immunoprecipitation studies identified another human protein, NFKBIL2.

NFKBIL2 shares several similarities with TONSOKU found in plants, hence it is called Tonsoku-like (TONSL). They both contain multiple protein interaction modules such as tetratricopeptide (TPR) repeats, ankyrin (ANK) repeats, leucin-rich repeats (LRR), and PB1 ubiqutin-like domain. The domain structure reveals a potential function of TONSL as a scaffolding protein. Such protein usually serves as bridge/scaffold by interacting with other proteins, meaning that TONSL likely interacts with other proteins and not only with MMS22L.[75, 76,77]

1.3.2.2 MMS22L-TONSL complex interacts with proteins involved in DNA replication

Mass spectrometry analyses revealed that MMS22L and TONSL interact with each other *in vivo.*[71,72] In accord, depletion of one protein negatively affected the protein levels of the other protein and *vice versa*, indicating that MMS22L-TONSL interaction is indispensable for mutual protein stability.[71]

Furthermore, MMS22L-TONSL interacts with RPA (RPA1 and RPA3 subunits), MCM (minichromosome maintenance protein complex) helicase (MCM2, -4, -6, -7) [71,72], histone chaperone FACT (facilitates chromatin transcription) complex and its subunits SUPT16H and SSRP1 [72], ASF1A and ASF1B histone chaperones [71], histones H2A, H2B [71], and H3.1 [71,78] (Figure 3A). Moreover, interaction between TONSL and MCM is facilitated via TRP repeats located at the N-terminus of TONSL with help of ANK repeats, whereas TONSL-MMS22L interaction is mediated via C-terminus of the TONSL where LRR repeats are located (Figure 3B).[73] Surprisingly, unlike yeast Mms22, no interaction with cullins or peptides derived from DDB1 (damage-specific DNA binding protein 1) was detected, suggesting that MMS22L is structured differently than the yeast homolog [71,72], and that human MMS22L will not be likely involved in protein ubiquitylation. The MMS22L-TONSL interaction data altogether suggest a strong link with DNA replication, where the MMS22L-TONSL complex might play a role.

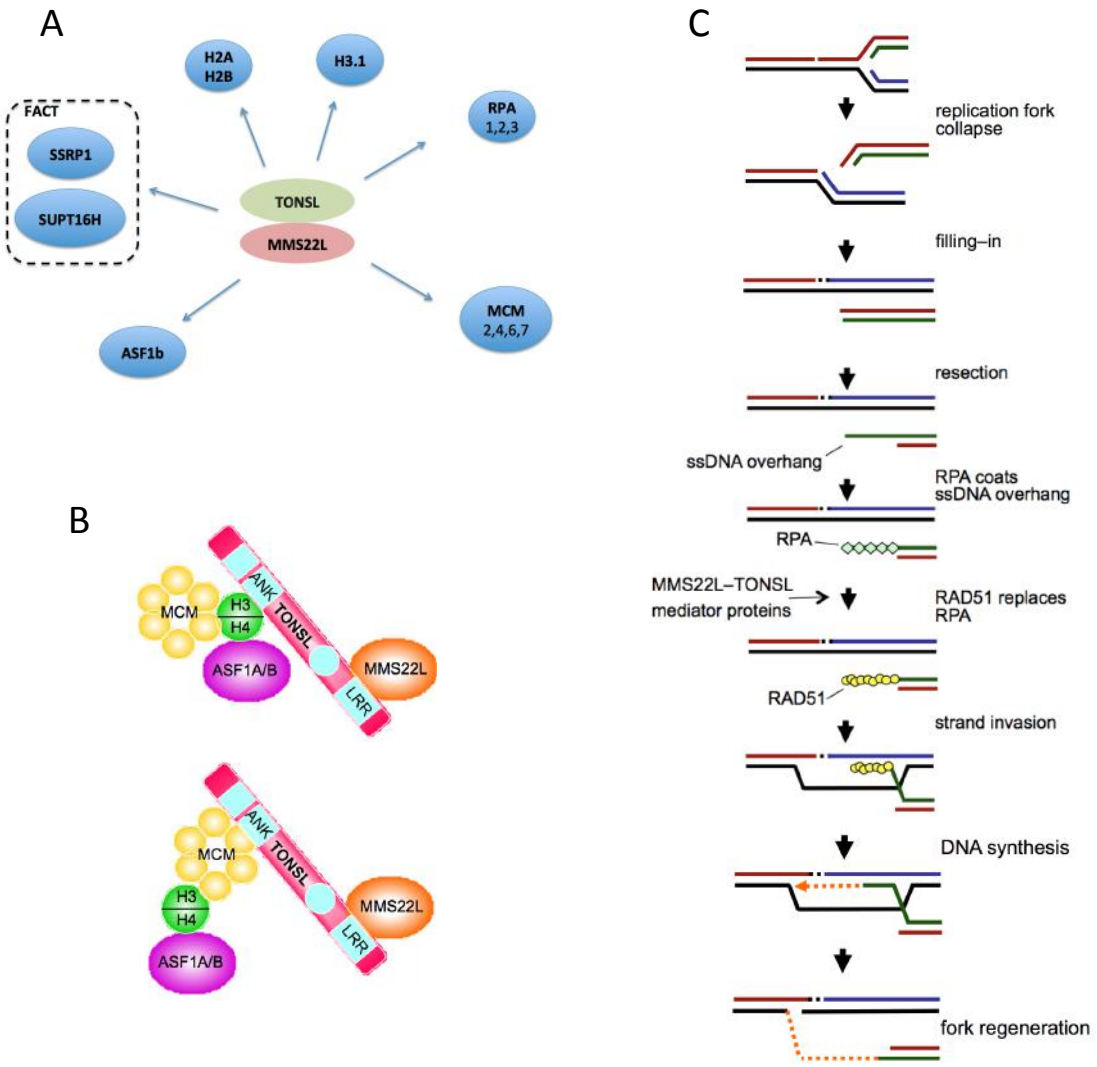


Figure 3: MMS22L-TONSL.

A- MMS22L-TONSL interacting partners. **B-** TONSL's interacting domains. **C -** Possible function of MMS22L-TONSL as a recombination mediator upon replication fork collapse (Adapted from Duro et al.[71], O'Donnel et al.[72], and Campos et al.[78]).

1.3.2.3 Possible function of MMS22L-TONSL in homologous recombination

Intriguing question was whether MMS22L-TONSL has a similar function as Mms22 in yeast. To investigate that, Duro et al. [71] depleted either MMS22L or TONSL, which led to elevated levels of DSBs, marked as 53BP1 and γ-H2AX subnuclear foci, even without introducing the DNA damaging agents.[71,72] CPT, HU, MMS, and ultraviolet [13] light can induce replication-associated ssDNA breaks. Treatment with

such DNA damaging agents resulted in accumulation of MMS22L and TONSL foci, hinting that MMS22L-TONSL repairs stalled (or broken) replication forks.[72]

In yeast, as mentioned above, Mms22 was required for HR repair upon replication fork stalling but not for HR induced by DSBs.

Such observation was further confirmed in the following experiment, where depletion of MMS22L or TONSL did not affect RPA foci formation upon CPT treatment, although these cells showed a recovery delay [71] [72], which points to the problem downstream of resection.

Ataxia telangiectasia mutated- and Rad3-related (ATR) signaling pathway predominantly responds to replication fork stalling and is recruited to ssDNA regions/lesions covered by RPA. ATR then signals the DNA damage by phophorylating CHK1 kinase to prevent a collapse of replication fork during S-phase of the cell cycle. [79,80,71,81,82] Depletion of either MMS22L or TONSL together with ATR sensitized cells to drugs hydroxyurea (HU), cisplatin and to ionizing radiation, which are known to interfere with progression of the replication fork machinery.[71] Moreover, MMS22L depletion led to the accumulation of cells in G2-phase of the cell cycle after the release of CPT.[72] Such event was further followed by the phosphorylation of the CHK1 kinase on serine 317 (S317) [72], demonstrating the checkpoint activation as a response to DNA damage and replication fork stalling.[82]

Indeed, depletion from MMS22L and TONSL resulted in a reduction of RAD51 foci and RAD51 foci focus formation was severely affected after treatment with CPT and IR [72], meaning that MMS22L-TONSL is factor acting upstream of RAD51 foci formation. Interestingly, depletion of MMS22L or TONSL resulted in elevated levels of 53BP1 foci similarly to RAD51 paralogs, such as RAD51C and XRCC3.

The final step was to investigate the involvement of MMS22L-TONSL in DR-GFP reporter assay, a common tool to study HR in mammalian cells. I-SceI endonuclease cuts at rare-sites and creates DSBs. Successfully repaired DSBs by HR can be easily detected by GFP signal.[83] Interestingly, downregulation of RAD51 resulted in about 96% reduction of GFP-positive cells and downregulation of both MMS22L and TONSL showed around 60% reduction in GFP-positive cells.[72]

This suggests, that MMS22L-TONSL complex facilitates RAD51 loading and acts as a recombination mediator in HR (Figure 3C), similarly as RAD51 paralogs or BRCA2 (Figure 4). Furthermore, depletion of neither MMS22L nor TONSL affected the

stability of RAD51 paralogs RAD51C and XRCC3 [71], and BRCA2[72], suggesting that MMS22L-TONSL promotes RAD51 loading in a distinct manner.

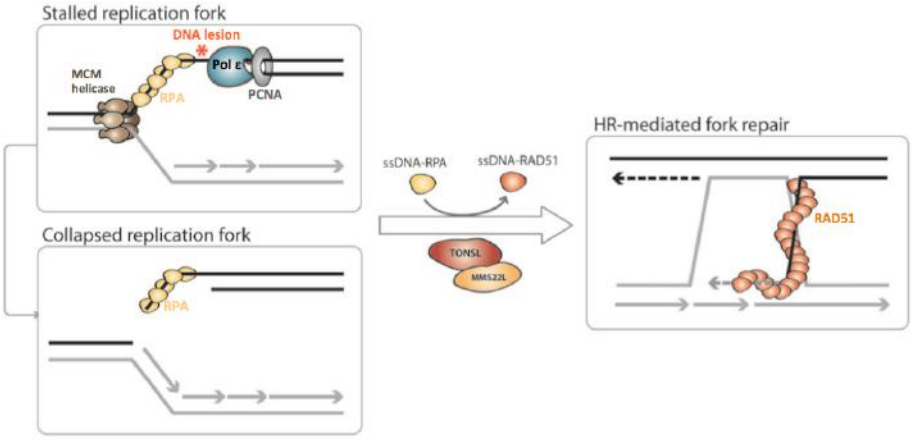


Figure 4: Potential HR-mediated fork repair of stalled and collapsed replication fork by MMS22L-TONSL (Adapted from Piwko *et al.* [84]).

1.4 LINKS BETWEEN RECOMBINATION, DISEASE AND CANCER

1.4.1 Hereditary versus sporadic cancers

Genome instability and mutations are hallmarks of cancer. Germline mutations in tumor suppressor genes lead to a hereditary threat of developing cancer.[85,86] On the other hand, 'sporadic cancers' arise from acquired mutations. Sporadic cancers can for example arise from a mutational or epigenetic inactivation. Example of epigenetic inactivation was described in case of BRCA1. Such event is a result of aberrant methylation of cytosine residues in a CpG island, which is located in the vicinity of the transcription start site of the gene. DNA is then more densely packaged by the histones and therefore may lead to a silencing of the transcription of the affected gene.[86] Both sporadic and hereditary cancers can be a result of aberrant mutations

that are present in genes coding for proteins, which play a crucial role in the repair mechanism of DNA, i.e. HR. Such mutations may lead to loss of function and hence faulty repair. Moreover, nonfunctional HR predisposes not only to cancer, but also can lead to infertility, birth defects, and negatively affects development.[65]

1.4.2 Defects in recombination genes lead to cancers, Fanconi anemia and developmental problems

Monoallelic mutations in *PALB2*, *BRCA1*, and *BRCA2* are linked to hereditary breast and ovarian cancers.[87,88] Similar outcome can be observed for RAD51 paralogs, where mutations in RAD51B, RAD51C, XRCC2, and XRCC3 lead to breast and ovarian cancers as well.[26,89] Moreover, mutation of PALB2 in one allele can also lead to a higher risk of development of pancreatic cancer.[26,90,89,91] Interestingly, women with mutations in PALB2 are predisposed eight to nine more times to develop cancer in comparison to unaffected population aged below 40 years.[91] Moreover, mutations in both alleles of *PALB2*, *RAD51C*, *BRCA1*, and *BRCA2* subsequently lead to Fanconi anemia, a rare childhood disease, which manifests various defects in development including bone marrow failure as well as increased susceptibility of the affected individuals to cancer.[26,92,93]

HR is a largely accurate DNA repair pathway, which is taking place mostly in the S-phase. The recombination machinery can accurately copy missing information from the homologous template and avoid unwanted introduction of a mutation into the DNA, hence it is a preferred repair pathway despite its rather slow progression. However, in the absence of HR, NHEJ or SSA pathways compensate for the repair defect, however such repair mechanisms are error-prone and may potentially introduce mutations or lead to deletions. Such scenario can arise especially when the critical factors in the repair machinery fail to perform their task.[86] For example, mutations interfering with BRCA2's ability to interact with crucial factors of HR such as RAD51 or RPA result in a failure to complete the repair. For example, mutation BRCA2 Y42C, which impairs RPA and BRCA2 interaction, predisposes affected individuals to cancer.[40]

BRCA2 contains two sequence motifs necessary for its transportation between the cytoplasm and nucleus. These are nuclear export sequence/signal (NES1; a region located between the amino acid sequence 1383-1393) [94,95] and nuclear localization sequence (NLS), which is located in the C-terminus of BRCA2.[95] NLS and NES are responsible for shuttling of BRCA2 from cytoplasm to nucleus across the nuclear envelope and *vice versa.*[94] Upon DNA damage, DSS1 masks NES, so that BRCA2 can be transported to the nucleus. Similarly, once DSS1 masks NES, BRCA2 can effectively transport RAD51 oligomers to the nucleus to the sites of DNA damage. In accord, $BRCA2^{D2723H}$ point mutation that impairs DSS1 binding to BRCA2 results in a failure to effectively transport RAD51 to the sites of damage [95] and is commonly associated with breast and ovarian cancer.[96]

Missing function of BRCA2 can also result in developmental problems. An example is the interaction between BRCA2 and DSS1, which is impaired in split hand/split foot (SHFM). As the name suggests, mutations in the DSS1 were originally identified in individuals affected by malformations of the limbs, hands and feet, resulting in a so called "lobster claw deformity" [97]. DSS1, an evolutionary conserved protein, binds the BRCA2 at amino acid residues 2472-2957.[97] Lack of such interaction leads to failure to complete the HR. Indeed, depletion of DSS1 resembles the phenotype of BRCA2-depleted cells, specifically in the hypersensitivity to DNA damaging agents. The importance of DSS1-BRCA2 interaction lies in the fact that DSS1 is crucial for delivery of BRCA2 (and RAD51) to the nucleus.[95]

1.4.3 Towards targeted therapy: exploiting synthetic lethal interactions

Knowing the exact molecular mechanism, which leads to the risk and/or development of cancer, is crucial in order to design appropriate cancer therapy, which can be targeted to the patient depending on the genetic footprint of the disease. Synthetic lethality is a phenomenon when two parallel pathways can lead to a similar biological outcome. Disabling one pathway does not have a great effect, however when both pathways are inactive, it results in a synthetic lethality.[98] An example for the exploitation of synthetic lethality is the treatment of BRCA2-deficient tumors with PARP inhibitors, marketed under the label Lynparza (olaparib).[99,100,98] Poly(ADP-ribose) polymerase (PARP1) is an enzyme that binds to

the sites of DNA damage. BRCA2-deficient cells/tumors were shown to be sensitive to PARP inhibitors, therefore this drug is used for killing the BRCA2-deficient tumors in triple negative breast/ovarian cancer patients.[100]

2 RESULTS

2.1 Summary of results

2.2 Preliminary results

2.2.1 DNA binding activity of MMS22L-TONSL and MMS22L alone

To determine whether MMS22L-TONSL binds DNA, we performed electrophoretic mobility shift assay (EMSA) with various radiolabelled DNA molecules. We found that MMS22L-TONSL did not bind to short oligonucleotides of 50 nt in length. The only exception was a 4-way junction, where low DNA binding was detected (Figure 5A). Instead, MMS22L-TONSL bound to longer DNA substrates (93 nt), albeit still with a low affinity. In general, we did not observe major differences between the various structures tested, although moderate preference for ssDNA (93 nt) and branched structures, such as 4-way junctions and replication fork substrates, was observed (Figure 5B). Nevertheless, the difference between the most and least preferred substrates was 2-3 fold. The relatively low concentration of our protein preparations did not allow us to establish K_d values. Still, based on the trend line we estimate DNA dissociation constants at approx. 50-120 nM range.

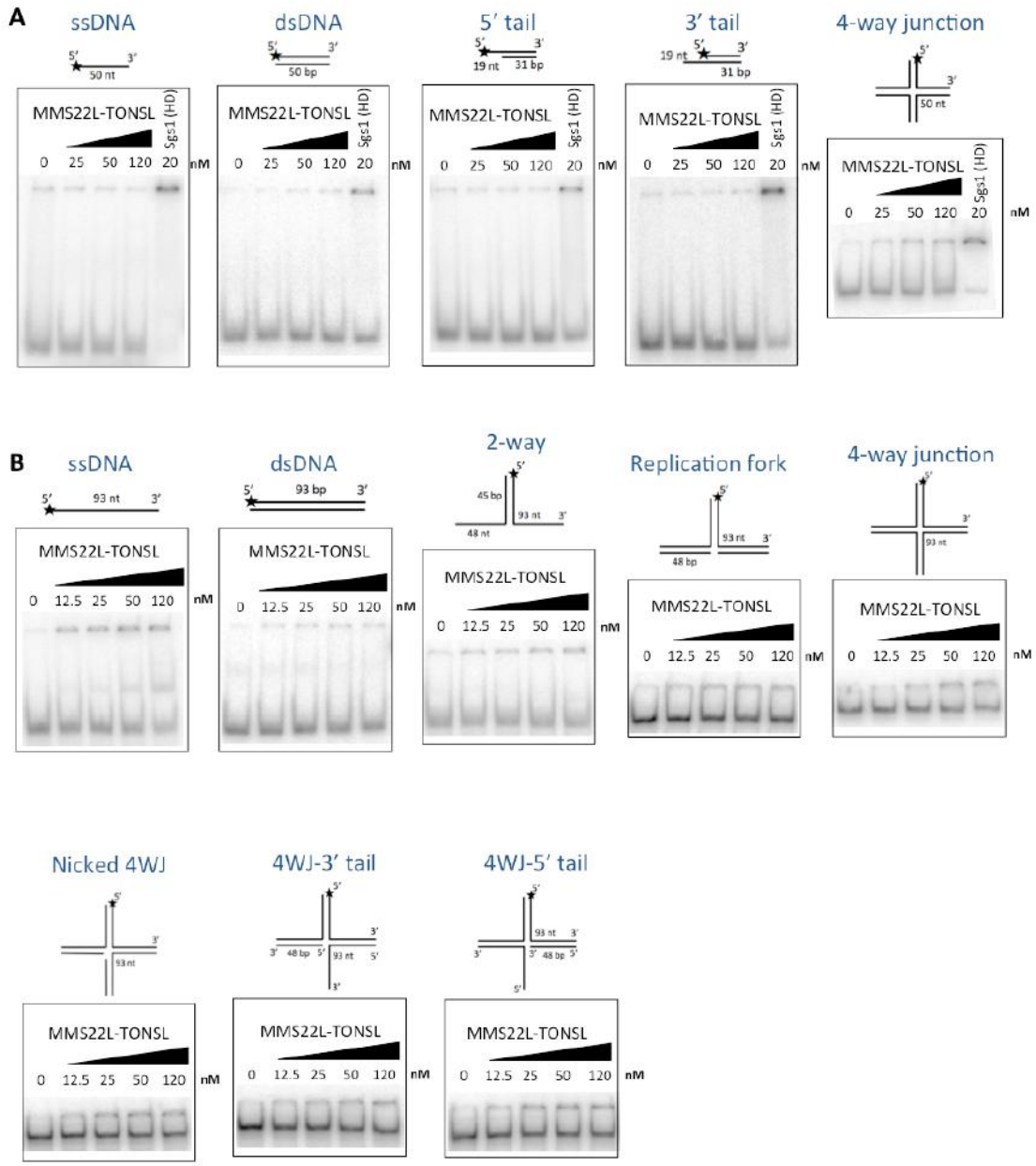


Figure 5: A- Binding of MMS22L-TONSL complex to 50nt long substrates. **B**- Binding of MMS22L-TONSL complex to 93 nt/bp substrates where 4WJ stands for '4-way junction'.

Next we tested MMS22L-TONSL binding to bacteriophage M13-based, 7250 nt/bp long DNA substrates. In contrast to the oligonucleotide-based structures, MMS22L-TONSL bound long DNA with a high affinity (Figure 6A). This suggested that the heterodimer oligomerizes upon DNA binding, and DNA-bound multimers bind DNA more strongly then monomers or small oligomers. We also tested the binding of

MMS22L protein alone, which showed dramatically lower DNA binding than the heterodimer. Thus, both subunits are required for stable DNA binding (Figure 6B and C).

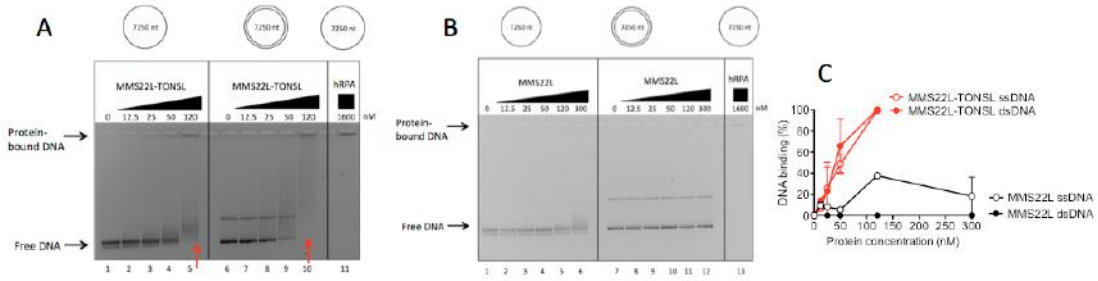


Figure 6: A, C- MMS22L-TONSL binds long stretches of ssDNA and dsDNA with high affinity. **B, C**- MMS22L alone does not bind long stretches of ssDNA and dsDNA. hRPA is a positive control.

2.2.2 The MMS22L-TONSL heterodimer directly promotes RAD51-dependent recombination upon replication stress

Wojciech Piwko, **Lucie J. Mlejnkova**, Karun Mutreja, Alexander Smirnov, Mia M. L. Brodersen, Diana Stafa, Ralph Zellweger, Andreas Sturzenegger, Pavel Janscak, Massimo Lopes, Matthias Peter, and Petr Cejka.

Manuscript.

I designed the biochemistry part of the research together with P.C. and performed the experiments. I analyzed the biochemistry data together with P.C. and wrote the manuscript together with P.C. and W.P.

The MMS22L-TONSL heterodimer directly promotes RAD51-dependent recombination upon replication stress

Wojciech Piwko[1,*,#], Lucie J. Mlejnkova[2,*], Karun Mutreja[2], Alexander Smirnov[2], Mia M. L. Brodersen[1], Diana Stafa[1], Ralph Zellweger[2], Andreas Sturzenegger[2], Pavel Janscak[2], Massimo Lopes[2], Matthias Peter[1,#], and Petr Cejka[2,#]

[1] Institute of Biochemistry, Department of Biology, ETH Zurich, 8093 Zurich, Switzerland

[2] Institute of Molecular Cancer Research, University of Zurich, 8057 Zurich, Switzerland

[#] to whom correspondence should be addressed:

cejka@imcr.uzh.ch

matthias.peter@bc.biol.ethz.ch

wojciech.piwko@bc.biol.ethz.ch

*these authors equally contributed to this study

Abstract

Homologous recombination is a key pathway that repairs DNA double-strand breaks (DSBs) and helps restarting stalled or collapsed replication forks. How HR supports replication upon genotoxic stress is not understood. Using *in vivo* and *in vitro* approaches, we show that MMS22L-TONSL localizes to replication forks under unperturbed conditions and its recruitment is increased during replication stress. MMS22L-TONSL heterodimers associate with replication protein A (RPA)-coated ssDNA, and MMS22L directly interacts with the strand exchange protein RAD51. Importantly, recombinant MMS22L-TONSL limits the assembly of RAD51 on dsDNA, which stimulates RAD51-ssDNA nucleoprotein filament formation and RAD51-dependent strand exchange activity *in vitro*. By specifically regulating RAD51 activity at uncoupled replication forks, MMS22L-TONSL promotes replication fork reversal *in vivo*, thereby stabilizing perturbed replication forks. These results provide unprecedented mechanistic insight into the molecular mechanism through which MMS22L-TONSL promotes genome stability.

Introduction

All cells must replicate their DNA before each round of cell division. In humans, genome duplication is critically dependent on homologous recombination (HR), which supports stalled replication forks or restores integrity of collapsed forks (San Filippo et al., 2008). Recombination mechanisms are largely studied in the context of DSB repair, yet endogenous two-ended DSBs are relatively rare and usually repaired by a non-homologous end-joining pathway in mammals. In contrast, HR is required for the completion of replication even in the absence of exogenous DNA damage, as recombination-deficient human cells undergo only a very limited number of DNA replication rounds, however, how recombination supports replication remains mostly undefined (Neelsen and Lopes, 2015; Petermann and Helleday, 2010). Furthermore, mutations in several regulatory HR genes including breast cancer susceptibility protein 2 (BRCA2) or RAD51C predispose affected individuals to cancer, demonstrating that aberrant HR results in genome instability that drives tumorigenesis (Meindl et al., 2010; Wooster et al., 1995).

A key player in the HR-mediated processes is the RAD51 recombinase, which assembles on ssDNA upon nuclease-mediated resection of the broken ends. RAD51 nucleoprotein filament then invades homologous DNA, which serves as the repair template. The assembly of RAD51 on ssDNA is regulated on multiple levels. First, resected ssDNA is rapidly coated and protected by RPA, which coordinates DNA damage signaling and repair. However, RPA also competes with RAD51 for binding to ssDNA. A group of proteins termed recombination mediators has the capacity to load RAD51 on RPA coated ssDNA. This overcomes the inhibitory effect of RPA on RAD51-mediated DNA strand exchange, and represents an important regulatory step that allows recombination on appropriate DNA substrates. These mediators include BRCA2 and the five RAD51 paralogues RAD51B/C/D and XRCC2/3 (San Filippo et al., 2008). These proteins help form the RAD51-ssDNA filament in the presence of RPA and additionally, in particular BRCA2, prevent assembly of RAD51 on dsDNA (Jensen et al., 2010; Liu et al., 2010; Sigurdsson et al., 2001b; Taylor et al., 2015; Thorslund et al., 2010). Together, the recombination mediators appear to

be the key to understanding how HR is fine-tuned, but their exact roles and interplay at DNA damage sites is still poorly understood.

The human MMS22L-TONSL complex and its budding yeast counterpart Mms22 play essential roles during DNA replication stress but their precise function remains elusive (Duro et al., 2010; Duro et al., 2008; Luke et al., 2006; O'Connell et al., 2010; O'Donnell et al., 2010; Piwko et al., 2010). Yeast Mms22 was proposed to be a putative substrate adaptor of the Rtt101 cullin E3 ubiquitin ligase functioning downstream of the histone H3-K56 acetylation pathway, which marks newly replicated DNA and is required for genome stability maintenance (Collins et al., 2007; Zaidi et al., 2008). Human cells expressing low levels of MMS22L or TONSL as well as yeast *mms22*Δ mutants are hypersensitive to drugs inducing DSBs at replication forks such as camptothecin (CPT) but not to ionizing radiation, suggesting specific defects in dealing with replication fork stalling or collapse (Duro et al., 2010; Duro et al., 2008; Luke et al., 2006; O'Connell et al., 2010; O'Donnell et al., 2010; Piwko et al., 2010). Indeed, cells lacking MMS22L were unable to restart replication forks stalled upon treatment with CPT (O'Donnell et al., 2010). Microscopy data demonstrated defects in RAD51 foci formation suggesting that MMS22L-TONSL promotes HR-mediated repair at DSBs and collapsed replication forks by regulating the assembly of RAD51 filaments, and thus was suggested to act as a recombination mediator (O'Donnell et al, 2010; Duro et al, 2010). However, evidence for a direct involvement of MMS22L-TONSL in this process has been missing. To this point, both yeast and human MMS22L were found in complexes with histones and/or chromatin remodeling factors, raising the possibility that they may act indirectly by regulating chromatin milieu or the mobility of broken DNA.

Here, we report that MMS22L-TONSL is present at nascent DNA under unperturbed conditions and its presence is enhanced at collapsed replication forks that had undergone extensive resection. MMS22L-TONSL is likely recruited to replication fork stall sites *via* a direct interaction with RPA-coated ssDNA. Its activity at stalled replication forks leads to replication fork protection by RAD51-dependent fork reversal. We demonstrate that purified MMS22L-TONSL directly binds the RAD51 recombinase, and favors the interaction of RAD51 with ssDNA, which promotes the formation of RAD51 presynaptic filaments and invasion of

homologous DNA. These data demonstrate a direct involvement of MMS22L-TONSL in RAD51-dependent recombination, and establish the heterodimer as a key factor that functions alongside RAD51 paralogues to promote recombination at stalled replication forks.

Results

MMS22L is recruited to stalled and collapsed replication forks

MMS22L-TONSL accumulates as nuclear foci upon induction of DSBs in human cells (Duro et al., 2010; O'Donnell et al., 2010). To determine whether MMS22L recruitment to DSBs is cell cycle regulated, we treated synchronized HeLa cells with the Topo II inhibitor etoposide. As expected, etoposide induced DSBs in all cell cycle phases, demonstrated by γH2AX foci accumulation (Nitiss, 2009), but only S-phase cells showed MMS22L foci (Figures 1A and 1B, and Figures S1A-C). When replication was inhibited by the polymerase-alpha inhibitor aphidicolin, the number of etoposide-induced MMS22L foci in S-phase cells was strongly reduced (Figures S1D and S1E), suggesting that MMS22L is not recruited to DNA breaks in general but likely acts specifically in the context of replication forks. Since *Saccharomyces cerevisiae mms22*Δ cells are severely impaired in HR specifically induced by drugs that inhibit DNA replication (Duro et al., 2008), the function of MMS22L in promoting replication-linked recombination appears conserved in evolution.

Key HR players such as RPA, recombination mediator BRCA2 and the strand exchange protein RAD51 accumulate as nuclear foci at DSBs (Raderschall et al., 1999; Tarsounas et al., 2004). To better understand the nature of replication stress-induced MMS22L foci, we investigated their co-localization and dependence on RPA, BRCA2 and RAD51 in human osteosarcoma U2OS cells. We observed that 47% BRCA2 and 43% RAD51 foci co-localized with MMS22L at stalled replication fork sites upon prolonged treatment with the Topo I inhibitor camptothecin (CPT) (Figures 1C and 1D), suggesting that MMS22L cooperates or functions alongside BRCA2 in the assembly of RAD51 presynaptic filaments at these sites. We found the greatest co-localization of MMS22L foci with those of RPA2 (~63%). Furthermore, the RPA2 and MMS22L foci signal intensity

correlated with each other upon various replication stress conditions (Figure S1F), suggesting that RPA may recruit MMS22L-TONSL to stalled replication forks. Indeed, disrupting RPA accumulation upon replication stress by the depletion of the DNA end resection factor CtIP strongly reduced the number of MMS22L foci (O'Donnell et al., 2010; Sartori et al., 2007), in contrast to depletion of downstream HR factors such as BRCA2 or RAD51 (Figures S1G-K).

MMS22L directly binds RPA-coated ssDNA

To understand the recruitment of the complex to replication stress sites at molecular level, we assessed whether MMS22L-TONSL directly interacts with RPA. We purified the full-length recombinant heterodimer and the MMS22L subunit from *Sf*9 cells (Figures S2A, and S2B). TONSL without MMS22L was rapidly degraded and could not be obtained (data not shown). MMS22L-TONSL bound recombinant RPA directly but with a weak affinity (Figures S2C and S2D). We found that the MMS22L-TONSL complex does not bind oligonucleotide-based ssDNA (Figure 2A, lanes 1-4). However, when we pre-bound RPA to ssDNA and introduced MMS22L-TONSL (Figures 2A and 2B, and Figures S2E-G, lanes 6-10), we observed a novel band likely corresponding to the ternary MMS22L-TONSL-RPA-ssDNA complex. To verify these findings, we incubated biotinylated ssDNA, RPA and his-tagged MMS22L-TONSL. We detected MMS22L-TONSL in streptavidin pulldowns only when RPA was present in the reaction (Figure 2C, compare lanes 3 and 5). Likewise, in NiNTA pulldowns, more RPA was obtained in reactions containing ssDNA (Figure 2C, compare lanes 6 and 7). Thus, MMS22L-TONSL forms a stable complex with RPA bound to ssDNA, which is dependent on the MMS22L subunit (Figures S2H-J). The RPA interaction likely facilitates MMS22L-TONSL recruitment to ssDNA present at or behind stalled and collapsed replication forks.

MMS22L-TONSL promotes RAD51 loading to stalled replication forks

We next wanted to understand if MMS22L-TONSL acts directly at DNA replication forks. To assess MMS22L-TONSL dynamics at replication forks, we employed the isolation of proteins on nascent DNA (iPOND) assay (Sirbu et al., 2011) based on affinity purification of DNA pulse-labeled during replication with the thymidine analog ethynyl-deoxyuridine (EdU). We found both MMS22L and

TONSL proteins at active forks under unperturbed conditions, and binding was strongly reduced on fully replicated chromatin (Figure 3A; compare lane 9 with lanes 10-12). Hydroxyurea (HU) addition to stalled forks led to a prolonged MMS22L-TONSL retention on the nascent DNA (Figure 3A; lanes 13-16), suggesting that the complex functions directly at stalled replication forks. We next stalled forks with a 30-minute CPT treatment, followed by drug wash out to monitor the recovery process over time. We observed a gradual RPA2 reduction along with loading of the RAD51 recombinase directly to forks (Figures 3B and 3C). MMS22L dynamics closely correlated with the exchange of RPA for RAD51. As the MMS22L-TONSL complex is required for RAD51 foci formation upon DNA damage (Duro et al., 2010; O'Donnell et al., 2010), we next tested if MMS22L-TONSL is required to recruit RAD51 to stalled forks. Using the CPT-recovery assay with cells depleted for MMS22L or TONSL, we observed a strong RAD51 reduction at CPT-stalled forks in the absence of either MMS22L (Figures 3D and 3E) or TONSL (Figures S3B and S3C). Moreover, ectopic expression of TONSL mutants defective in MMS22L interaction reduced RAD51 recruitment to stalled forks (Figures S3A-C), indicating that an association of both components of the complex is required to promote RAD51 loading. These data strongly suggest that MMS22L-TONSL is involved in an efficient recruitment and/or maintenance of the RAD51 recombinase directly on DNA replication forks immediately upon stalling.

MMS22L primes replication fork reversal

Stalled replication forks may either collapse resulting in single-ended DSBs or be transiently stabilized through fork reversal (Neelsen and Lopes, 2015). Micromolar CPT concentrations that induce replication associated DSBs resulted in early MMS22L foci (Figures S4A-C). Instead, nanomolar CPT or HU that stall replication forks, but do not immediately lead to DSBs (Petermann et al., 2010; Ray Chaudhuri et al., 2012), did not result in detectable MMS22L foci at similar time points (Figures S4A-C). MMS22L foci number and intensity were strongly increased in HU-treated cells upon ATR kinase inhibition, which leaves stalled forks unprotected from collapse (Cimprich and Cortez, 2008) (Figures S4D and S4E). These observations suggest that MMS22L foci detectable by conventional

microscopy mainly form upon DNA breakage, and thus implicate MMS22L-TONSL in the recombination-mediated repair of collapsed replication forks. Thus, as described for RAD51 (Petermann et al., 2010), MMS22L-TONSL is recruited to stalled forks early to support replication fork restart (Figure 3A) without creating visible foci (Figures S4A-C), while it forms bright foci in the HR process that repairs collapsed forks (Figure 1).

Recently, RAD51 activity at CPT-stalled forks has been linked to fork reversal in human cells (Zellweger et al., 2015). This allows transient fork pausing and was proposed as a mechanism to avoid collision with lesions or gaps on template DNA, thus preventing fork collapse (Neelsen and Lopes, 2015). Electron microscopy analysis of *in vivo* psoralen-crosslinked replication intermediates from MMS22L-depleted U2OS cells revealed a marked reduction (~60%) in the frequency of CPT-induced reversed forks (Figures 4A-C, and Figure S4F), similar to defects recently reported upon depletion of RAD51 (Zellweger et al., 2015). Moreover, forks in CPT-treated MMS22L-depleted cells displayed significantly longer ssDNA regions (Figures 4B and 4D), strongly suggesting that MMS22L-TONSL is recruited to extended ssDNA gaps on template DNA, thereby assisting fork reversal.

MMS22L-TONSL facilitates RAD51-dependent DNA strand exchange

To determine whether MMS22L-TONSL directly promotes RAD51-dependent recombination, we employed *in vitro* DNA strand exchange assays (Figure 5A). While a ssDNA-bound RAD51 filament is the active species capable to search and invade homologous DNA, RAD51 bound to dsDNA is inhibitory for recombination (Benson et al., 1994; Sigurdsson et al., 2001a). Reconstituted strand exchange reactions thus show a very narrow RAD51 concentration optimum. After RAD51 fully saturates ssDNA (~130 nM, Figure 5A and 5B), further increase of RAD51 concentration results in its binding to dsDNA, which inhibits DNA strand exchange (Benson et al., 1994). When supplemented with MMS22L-TONSL (75 nM), the reaction optimum was reached at a similar RAD51 concentration (~130 nM, Figure 5C), but strikingly the inhibition at higher RAD51 concentrations was partially alleviated (Figures 5C and 5D). A similar behavior was observed in strand exchange reactions with both 3′ and 5′ tailed as

well as gapped DNA (Figures S5A-C), showing that the stimulatory effect of MMS22L-TONSL is not limited to a specific DNA structure and does not require a terminus, in keeping with the fork reversal model described above. All reactions were ATP-dependent and required RAD51 (Figures S5D and S5E), implying that MMS22L-TONSL is not capable of DNA strand exchange alone, but functions *via* promoting RAD51 (Sung, 1994). We then used a high RAD51 concentration (270 nM) and titrated MMS22L-TONSL or MMS22L into the reactions. The heterodimer promoted strand exchange in a concentration-dependent manner showing saturation at ~50 nM (Figure 5E), while the MMS22L subunit alone was largely inactive (Figure 5E and Figures S5F and S5G). Therefore, both MMS22L and TONSL are required to stimulate DNA strand exchange.

MMS22L reduces affinity of RAD51 to dsDNA

We found that MMS22L-TONSL directly interacts with RAD51 (Figure 5F), with up to two RAD51 molecules tightly bound by the heterodimer (affinity in the nanomolar range, Figure 5G and Figure S5J). MMS22L-TONSL binds RAD51 with similar affinities as the MMS22L subunit (Figures S5J-M). However, MMS22L alone did not promote DNA strand exchange, and both MMS22L and TONSL were required (Figure 5E). Therefore, the stimulatory effect of MMS22L-TONSL is not a simple result of chelating and reducing the free RAD51 pool. Rather, these results infer that MMS22L-TONSL might function similar to the BRCA2 tumor suppressor by reducing RAD51 affinity for dsDNA (Jensen et al., 2010; Liu et al., 2010; Thorslund et al., 2010). To test this directly, we performed electrophoretic mobility shift assays with RAD51 and the MMS22L-TONSL heterodimer. Indeed, MMS22L-TONSL reduced RAD51 binding to dsDNA in a concentration-dependent manner, but had no effect on RAD51 binding to ssDNA (Figure 5H). This activity was specific to the heterodimer, as the MMS22L subunit did not alter binding to either ssDNA or dsDNA (Figures S5H and S5I). Finally we used competition pulldown assays to test whether the reduction of RAD51 binding capacity to dsDNA might promote its affinity for ssDNA (Figures 5I-K). As expected (Jensen et al., 2010), increasing the amount of dsDNA resulted in correspondingly less RAD51 bound to ssDNA. Conversely, titration of MMS22L-TONSL into reactions containing both ssDNA and dsDNA resulted in

more RAD51 bound to ssDNA, an effect not observed with the MMS22L subunit (Figure 5K). These results collectively demonstrate that the MMS22L-TONSL heterodimer reduces the RAD51 capacity to bind dsDNA, which likely facilitates RAD51 assembly on stretches of ssDNA at perturbed replication forks, and thereby promotes HR-mediated events.

MMS22L-TONSL and RAD51 paralogues cooperate at stalled replication forks

Recombination downstream of DNA end resection is stimulated by the BRCA2 tumor suppressor (Jensen et al., 2010; Liu et al., 2010; Thorslund et al., 2010), which relieves the inhibitory effect of RAD51 binding to dsDNA, and also loads RAD51 onto RPA-coated ssDNA. Both our *in vivo* and *in vitro* experiments show that although MMS22L-TONSL is recruited to RPA-coated ssDNA and reduces the ability of RAD51 to bind dsDNA, MMS22L-TONSL alone is unable to promote recombination on RPA-coated ssDNA (Figure S5N). Interestingly, however, this missing recombination mediator function was demonstrated for the RAD51B-RAD51C complex (Sigurdsson et al., 2001b), suggesting that MMS22L-TONSL likely functions in recombination alongside with BRCA2 and/or the RAD51 paralogue complexes. To test this hypothesis, we investigated the effect of MMS22L/TONSL depletion alone or in combination with other recombination mediators on the formation of RAD51 foci upon prolonged treatment with CPT. As expected (Jensen et al., 2013; Qing et al., 2011), BRCA2 depletion had a much stronger effect than depletion of MMS22L, almost completely preventing CPT-induced RAD51 foci formation (96% vs 41% foci number reduction, respectively; Figures S6A and S6B), while knockdown of RAD51B, RAD51C or XRCC3 alone had partial effects (34%, 56% or 54% reduction, respectively; Figures 6A and 6B). Importantly, co-depletion of MMS22L or TONSL with either of RAD51 paralogues led to synergistic effects, particularly strong in combination with RAD51C or XRCC3 (~90% reduction; U2OS cells: Figures 6A and 6B and Figures S6C-F; HeLa cells: Figure S6G). These results strongly suggest that MMS22L-TONSL functionally interacts with other recombination mediators in RAD51 nucleoprotein filaments assembly at the sites of replication fork restoration.

Discussion

The integrity and repair of stalled or collapsed replication forks is guaranteed by homologous recombination through yet largely undefined mechanisms. Previous reports indicated that the human MMS22L-TONSL heterodimer plays a crucial role in this process. Here we demonstrate that MMS22L-TONSL directly promotes recombination upon replication stress. We show that MMS22L-TONSL is present at nascent DNA under unperturbed conditions. Upon fork stalling, MMS22L-TONSL promotes replication fork reversal, which helps preventing replication fork collapse (Figure 7A). Additionally, the complex is enriched at collapsed replication forks that had undergone extensive resection (Figure 7B). Our data suggest that MMS22L-TONSL at these sites directly promotes invasion of homologous DNA by promoting the activity of the RAD51 strand exchange protein, leading to replication fork restoration or repair.

Why is MMS22L-TONSL particularly important for the repair of replication forks? A unique feature of MMS22L is its ability to preferentially bind RPA-coated ssDNA. This interaction may allow MMS22L-TONSL to dynamically associate with ssDNA produced by DNA unwinding during DNA replication or with extended ssDNA tracts that accumulate upon replication fork stalling. Indeed, our iPOND experiments demonstrate that MMS22L-TONSL dynamics at replication forks correlates well with that of the RPA2 protein (Figure 3). The interaction with RPA-coated ssDNA also likely enables the enrichment of MMS22L at collapsed replication forks that had undergone extensive resection. In addition, TONSL has been shown to directly bind histones (Duro et al., 2010) and the specific interaction with histone H4 unmethylated at K20 seems to contribute to the recruitment of the MMS22L-TONSL complex to newly replicated DNA, when a homologous template is available for repair (A. Groth, personal communication). This interaction could also facilitate MMS22L-TONSL access to postreplicative ssDNA gaps. Thus, the MMS22L-TONSL complex seems ideally equipped to participate in the repair of stalled and collapsed replication forks that accumulate stretches of ssDNA. Since the basal fork localization of the complex precedes the activation of the S phase checkpoint, MMS22L-TONSL may

be among the first factors helping the nucleation of RAD51 filaments not only upon replication stress but also upon endogenous fork pausing conditions. In support of this, the depletion of human MMS22L or TONSL, as well as the deletion of yeast *MMS22* gene led to accumulation of spontaneous DNA breaks during unperturbed S phase (Duro et al., 2010; Luke et al., 2006; O'Connell et al., 2010; O'Donnell et al., 2010; Piwko et al., 2010). This is in agreement with our observations that MMS22L-TONSL promotes fork reversal (Figure 4), which emerged as an important global response to replication stress and can occur in the absence of DNA breaks (Neelsen and Lopes, 2015; Ray Chaudhuri et al., 2012; Zellweger et al., 2015). Given that MMS22L- or TONSL-depleted cells are hypersensitive to low nanomolar CPT treatments that induce robust fork reversal but not to ionizing radiation causing DSBs (Duro et al., 2010), it is tempting to speculate that RAD51-dependent fork reversal represents the essential activity of the MMS22L-TONSL complex. Importantly - by MMS22L-TONSL-mediated nucleation of RAD51 on extended RPA-coated ssDNA - cells appear to simultaneously prevent DSB formation (by assisting fork reversal) and to repair the subset of forks that unavoidably collapsed (Fig. 7).

How does MMS22L-TONSL promote recombination? We demonstrate that recombinant MMS22L directly binds the RAD51 recombinase and provide evidence for a direct involvement of MMS22L-TONSL in promoting DNA strand exchange by RAD51 (Figure 5). The activity of the complex favors interaction of RAD51 with ssDNA over dsDNA, therefore promoting the formation of RAD51-ssDNA presynaptic nucleoprotein filaments. This is a prerequisite for invasion of homologous DNA and further steps in the HR pathway. MMS22L-TONSL acts as a mediator in a similar way as previously demonstrated for human BRCA2 (Jensen et al., 2010). However, unlike BRCA2, MMS22L-TONSL cannot catalyze the assembly of RAD51 filaments on ssDNA coated with RPA, even though it is directly recruited to such substrates *in vitro* and *in vivo*. This suggests that MMS22L-TONSL does not regulate RAD51 on its own, but RAD51-dependent repair likely requires the interplay of multiple recombination mediators. BRCA2 and MMS22L colocalize at foci upon CPT-induced replication stress. Additionally, we demonstrate that simultaneous depletion of MMS22L-TONSL and RAD51 paralogues (in particular RAD51B, RAD51C and XRCC3) leads to a dramatic

decrease of replication stress-induced RAD51 foci (Figure 6). This suggests that the MMS22L-TONSL heterodimer functionally interacts with these recombination mediators in the efficient assembly of RAD51 filaments upon replication stress *in vivo*. RAD51 paralogues may complement MMS22L-TONSL activity by alleviating the inhibitory effect of RPA, as such activity was previously demonstrated for both the RAD51B-RAD51C complex (Sigurdsson et al., 2001b) as well as BRCA2 (Jensen et al., 2010). Recently, it has been demonstrated that yeast RAD51 paralogues Rad55-Rad57 protect Rad51-ssDNA filaments from disruption by the antirecombinase Srs2 (Liu et al., 2011), while the RAD51 paralogues RFS-1/RIP-1 from *Cenorabditis elegans* are able to directly remodel and stabilize open conformation of ssDNA-RAD-51 filaments (Taylor et al., 2015). Human RAD51 paralogues have functions that partially overlap with those of BRCA2 (Sigurdsson et al., 2001b). Our data establish that MMS22L-TONSL belongs to the growing group of factors that regulate recombination, however in contrast to the other mediators its function seems to be highly specific to stalled or collapsed replication forks. Elucidating the precise relationship between the various recombination mediators and MMS22L-TONSL at stalled replication forks represents an exciting future challenge, in particular as the interplay between BRCA2 and the RAD51 paralogs remains unclear after many years of research.

The strong connections between defects in HR and tumorigenesis suggest that mutations in the *MMS22L* and *TONSL* genes might be associated with cancer development. It has been demonstrated that defects in recombination mediators including BRCA2, RAD51B and RAD51C frequently associate with breast and ovarian cancers (Golmard et al., 2013; Meindl et al., 2010; Wooster et al., 1995). Cancer genome sequencing projects identified several somatic mutations in the *MMS22L* and *TONSL* genes (Forbes et al., 2015), however, none of these alterations has been confirmed as tumorigenic. Here, we demonstrate that the two point mutations in TONSL (E1089K D1104N) that were identified in a genome of an ovarian cancer patient could potentially be tumorigenic (Forbes et al., 2015), as they abolish the interaction with MMS22L and negatively influence the recruitment of RAD51 to distressed replication forks (Figure S3). Moreover, we identified several additional cancer patient-derived point mutations in both

MMS22L and *TONSL* that affect protein stability (data not shown), suggesting that misregulation of the MMS22L-TONSL dependent processes might be involved in carcinogenesis. We therefore speculate that genome instability in some types of ovarian and possibly breast cancers might be promoted by *MMS22L* or *TONSL* mutations.

Acknowledgements

We thank A. Groth for communicating unpublished results; S. Taylor for the FRT-TetR-HeLa cell line; J. Matos for providing anti-RAD51C antibody; O. Oros, J. Tilma, S. S. Lee, F. van Drogen and the ETH ScopeM microscopy facility for technical support; A. Smith for critical reading of the manuscript and members of the Peter and Cejka labs for helpful discussions and comments on the manuscript. This work was supported by Oncosuisse, an ERC senior award, the Swiss National Science Foundation (SNF) and the ETH Zurich.

Author contributions

W.P., D.S. and M.L.B. performed cell biology experiments, and W.P. and M.P. analyzed the data; A.S. prepared expression constructs, L.J.M. performed biochemistry experiments, and L.J.M. and P.C. analyzed the data. K.M. and R.Z. performed EM analysis of reversed forks; A.S. and P.J. provided the polyclonal anti-RAD51 antibody and the modified pTXB3 expression vector; W.P., L.J.M., K.M., M.L., M.P. and P.C. participated in the experimental design; W.P., L.J.M., M.P. and P.C. wrote the paper.

Figure Legends

Figure 1. MMS22L foci form during S phase of the cell cycle and co-localize with HR proteins.

(A) MMS22L foci form at etoposide (ETP)-induced DSBs during S-phase of the cell cycle. Scale bar: 5 μm.

(B) Quantification of a representative experiment from (A) ($n = 2$; $n_{nuclei} \geq 60$).

(C) MMS22L foci co-localize with RPA2, BRCA2 and RAD51 foci in U2OS cells. Scale bar: 5 μm.

(D) Quantification of a representative experiment ($n = 3$) from (C) shows percentage of γH2AX, RPA2, BRCA2 and RAD51 foci co-localizing with MMS22L foci ($n_{nuclei} = 30$; nuclei with ≥ 25 MMS22L foci were analyzed).

(B, D) Boxplots represent distributions per nucleus; boxes indicate the 25–75 percentile and whiskers the 10–90 percentile. Horizontal lines mark the medians. Statistical analysis: Mann–Whitney *U* test; *** $P \leq 0.0001$; *ns*, not significant.

Figure 2. MMS22L-TONSL directly interacts with RPA-coated ssDNA.

(A) Electrophoretic mobility shift assay. MMS22L-TONSL does not bind ssDNA without RPA (lanes 1-4), but binds to RPA-coated ssDNA (lanes 6-10).

(B) Quantification of data as in (A), $n = 2$; error bars, SEM.

(C) Streptavidin and NiNTA pulldown of his-tagged MMS22L-TONSL, RPA and biotinylated ssDNA. The proteins in the input (lanes 1 and 2) and eluates (lanes 3-9) were detected by western blotting.

Figure 3. MMS22L-TONSL localizes to active and stalled replication forks.

(A-E) iPOND assays.

(A) MMS22L-TONSL is present at active and stalled replication forks. Antibodies used for immunoblots are indicated.

(B, C) MMS22L dynamics at stalled replication forks. Quantification shows average relative abundance of indicated proteins from immunoblots, $n = 3$; error bars, SEM.

(D, E) MMS22L depletion leads to reduced RAD51 levels at stalled replication forks. Quantification shows average relative RAD51 abundance, $n = 2$; error bars,

SEM.

Figure 4. MMS22L is required to promote reversal of uncoupled replication forks.

(A, B) Electron micrographs of a representative reversed fork and a fork with extended ssDNA region at the junction (arrow). P, parental duplex; R, regressed arm.

(C) Frequency of reversed replication forks detected by EM. Similar results were obtained in one independent experiment; $n = 2$.

(D) Distribution of ssDNA length at replication forks isolated from U2OS cells transfected with the indicated siRNAs and treated with 25 nM CPT for 1h. In brackets, the total number of analyzed molecules is given. Statistical analysis according to Mann–Whitney *U* test; ** $P \leq 0.01$, $n = 2$.

Figure 5. MMS22L-TONSL promotes DNA strand exchange by reducing RAD51 binding to dsDNA.

(A) Schematic representation of the DNA strand exchange assay with ssDNA.

(B-D) Strand exchange assay with ssDNA and RAD51 with (C) or without (B) the addition of MMS22L-TONSL (75 nM).

(D) Quantification shows averages, $n = 2$; error bars, SEM.

(E) Average of quantified strand exchange assays with ssDNA, RAD51 (270 nM) and varying concentrations of MMS22L-TONSL or MMS22L. $n = 2$; error bars, SEM.

(F) Amylose pulldown of MBP (control) or MBP-tagged MMS22L-TONSL and RAD51. Eluate proteins were detected by silver staining. MT: MBP-MMS22L-GST-TONSL.

(G) Quantitative analysis of RAD51 binding to MMS22L-TONSL such as shown in Figure S5J. Averages shown, $n = 2$; error bars, SEM.

(H) Quantification of electrophoretic mobility shift assays with dsDNA or ssDNA, RAD51 and MMS22L-TONSL or MMS22L. Averages shown, $n = 2$; error bars, SEM.

(I) A scheme of biotin pulldown assay.

(J) A representative western blot experiment showing reduced RAD51 binding to ssDNA (1 nM) upon addition of dsDNA.

(K) A representative western blot experiment showing MMS22L-TONSL and MMS22L titration into reactions containing RAD51, ssDNA (1 nM) and dsDNA (10 nM), as indicated.

Figure 6. MMS22L acts synergistically with RAD51 paralogues to promote RAD51 functions upon replication stress.

(A) Representative images showing the effect of co-depletion of MMS22L with RAD51B, RAD51C and XRCC3 on the number of RAD51 foci formed upon CPT treatment (50 nM, 18 hours) in U2OS cells. Scale bar: 20 μm.

(B) Quantification of RAD51 to γH2AX foci ratios per nucleus from a representative experiment (n = 4). Nuclei with ≥ 25 γH2AX foci (marker of replication stress sites) were analyzed; $n_{nuclei} \geq 145$. Boxplots represent distributions of foci per nucleus; boxes indicate the 25–75 percentile and whiskers the 10–90 percentile. Horizontal lines mark the medians. Statistical analysis: Mann–Whitney U test; *** $P \leq 0.0001$.

Figure 7. MMS22L-TONSL functions during replication stress.

Model of MMS22L-TONSL function at stalled (A) and broken (B) replication forks. MMS22L-TONSL is present near active replication forks, and promotes fork reversal upon stalling. In contrast, broken forks are resected by a CtIP-dependent mechanism, and repaired by break-induced replication. Our results suggest that MMS22L-TONSL heterodimer controls fork metabolism by reducing the capacity of RAD51 to bind dsDNA, which together with RAD51 paralogues and/or BRCA2 facilitates the formation of RAD51 filaments on stretches of ssDNA. This promotes invasion of homologous DNA, leading to replication fork restoration.

Supplemental Data Figure Legends

Figure S1. MMS22L foci formation during replication stress depends on RPA recruitment to resected ssDNA.

(A) Immunostaining of HeLa cells depleted for MMS22L confirming specificity of the anti-MMS22L antibody. Scale bar: 20 μm.

(B) Workflow and image analysis of the CellProfiler-based automated image processing pipeline used in this study to quantify nuclear speckles from immunofluorescence microscopy images. Scale bar: 5 μm.

(C) Related to the experiment in Figures 1A and 1B; DNA content analysis of synchronized HeLa cells using propidium iodide staining and flow cytometry.

(D) ETP-induced MMS22L foci formation depends on active replication forks in S-phase synchronized HeLa cells. APH, polymerase-alpha inhibitor aphidicolin. Scale bar: 5 μm.

(E) Quantification of a representative experiment from (D); $n = 2$. Graph represents distribution of foci numbers per nuclei, $n_{nuclei} \geq 177$.

(F) Signal intensities of colocalizing MMS22L and RPA2 foci correlate. The intensity of foci from U2OS cells treated with the indicated replication stress agents were quantified using CellProfiler. *r*, Pearson correlation coefficient; $n = 3$.

(G) Efficient MMS22L foci formation depends on CtIP-mediated DNA end resection in U2OS cells treated with indicated replication stress agents. Cells in silico enriched in S-phase based on increased nuclear RPA2 signal were analyzed. Boxplots show quantifications of MMS22L and RPA2 foci from a representative experiment, $n = 3$.

(H) qRT–PCR analysis demonstrating *CtIP* depletion efficiency; error bars, SEM.

(I) MMS22L foci formation does not depend on the presence of BRCA2 or RAD51 in U2OS cells. Nuclei with ≥ 25 RPA2 foci (marker of replication stress sites) were analyzed. Scale bar: 5 μm.

(J) Quantification of (I).

(K) Cells treated with the indicated siRNAs were collected and analyzed for BRCA2 and RAD51 by western blotting. Tubulin serves as loading control.

(E) (G) (J) Boxes indicate the 25–75 percentile and whiskers the 10–90 percentile. Horizontal lines mark the medians. Statistical analysis: Mann–

Whitney *U* test; *** $P \leq 0.0001$; *ns*, not significant.

Figure S2. Recombinant MMS22L-TONSL interacts with RPA-coated ssDNA.
(A) (B) Schematic representation of the recombinant constructs and polyacrylamide gels showing samples from MMS22L-TONSL (A) and MMS22L (B) purifications. MBP, maltose-binding protein; GST, glutathione S-transferase; PP, PreScission protease.
(C) (D) Amylose pulldown of MBP-tagged MMS22L-TONSL (MBP-MT) and RPA. The proteins in the eluates were analyzed by western blotting (C) or silver staining (D). RPA interacts with the MMS22L-TONSL heterodimer.
(E-G) Electrophoretic mobility shift assays with 5' tailed (E), 3' tailed DNA (F) and MMS22L-TONSL and/or RPA. The assays were quantified and shown as average (G), $n = 2$; error bars, SEM. MMS22L-TONSL only binds these DNA substrates if coated with RPA.
(H) (I) Electrophoretic mobility shift assay with MMS22L and ssDNA, with or without RPA. Panel (H) shows a representative experiment. The assays were quantified and shown as average (I), $n = 2$; error bars, SEM.
(J) Streptavidin pulldown of MMS22L, RPA and biotinylated ssDNA. The proteins in the input (lanes 1 and 2) and eluate (lanes 3-5) were detected by western blotting.

Figure S3. MMS22L-TONSL regulates RAD51 recruitment at stalled replication forks.
(A) Schematic representation of TONSL, TONSL^{LRR*}: E1089K and D1104N mutations within the LRR-domain identified in ovarian cancer patient (COSMIC sample OCC06PT(Forbes et al., 2015)) and TONSL$^{\Delta LRR}$: C-terminal TONSL truncation deleting the LRR domain; LRR, leucine-rich repeats. Co-immunoprecipitation of HSS-tagged TONSL, TONSL^{LRR*} and TONSL$^{\Delta LRR}$ with MMS22L and the histone-chaperone ASF1A (an interactor of the N-terminal region of TONSL (Duro et al., 2010; O'Donnell et al., 2010)).
(B) (C) The presence of the indicated proteins at replication forks was analyzed by the iPOND-CPT-release assay in cells expressing either HSS-tagged wild-type TONSL, or TONSL^{LRR*} or TONSL$^{\Delta LRR}$ mutants and depleted for endogenous TONSL. The experiments were quantified and shown as average (C) $n = 2$ except for LRR*

n = 1; error bars, SEM. TONSL and its interaction with MMS22L are required for normal levels of RAD51 at stalled replication forks.

Figure S4. MMS22L forms foci upon replication fork collapse.

(A) Representative microscopy images and boxplots of endogenous MMS22L (B) and RPA2 (C) foci distribution in U2OS cells treated with the indicated replication stress inducing agents. Cells *in silico* enriched in S-phase based on increased nuclear RPA2 signal were analyzed. Quantification of a representative experiment is shown; *n* = 3. Scale bar: 5 μm.

(D) (E) Nuclear MMS22L and RPA foci accumulate rapidly in U2OS cells upon collapse of HU-stalled replication forks by treatment with the ATR inhibitor (ATRi). Cells in silico enriched in S-phase based on increased nuclear RPA2 signal were analyzed. Scale bar: 5 μm. Graphs represent distribution of foci numbers per nucleus; boxes indicate the 25–75 percentile and whiskers the 10–90 percentile. Horizontal lines mark the medians. Statistical analysis: Mann–Whitney *U* test; *** $P \leq 0.0001$.

(F) Western blot analysis of cell extracts from cells used in a representative EM-experiment described in Figure 4. GAPDH serves as loading control.

Figure S5. MMS22L-TONSL promotes DNA strand exchange by RAD51.

(A-C) Quantification of DNA strand exchange with RAD51, with or without MMS22L-TONSL (75 nM) using a 3' tailed (A), 5' tailed (B) or gapped (C) DNA substrate. Averages are shown, *n* = 2; error bars, SEM.

(D) A representative DNA strand exchange experiment with ssDNA and RAD51, with or without ATP. The reaction requires ATP.

(E) A representative DNA strand exchange experiment with ssDNA, RAD51, MMS22L-TONSL (MT), with or without ATP, as indicated. The MMS22L-TONSL does not catalyze strand exchange on its own, but functions *via* stimulating RAD51 in ATP-dependent manner.

(F) (G) Representative DNA strand exchange experiments with a fixed concentration of RAD51 (270 nM) and varying concentrations of MMS22L-TONSL (F) or MMS22L (G).

(H) (I) Representative electrophoretic mobility shift assay experiments with MMS22L-TONSL, MMS22L, RAD51 and either dsDNA (H) or ssDNA (I). MMS22L-

TONSL reduces RAD51 binding to dsDNA, but does not affect binding of RAD51 to ssDNA.

(J-M) MMS22L-TONSL physically interacts with RAD51.

(J) Quantitative analysis of RAD51 binding to MMS22L-TONSL. Proteins were mixed in the indicated ratio, the MMS22L-TONSL complex was retained on amylose beads, and proteins in the eluate were analyzed by silver staining. Lane 6, the same amount of RAD51 was used as in lane 5, but no MMS22L-TONSL (MT) was added to control for non-specific binding of RAD51 to the resin.

(K) A NiNTA pulldown of his-tagged MMS22L and RAD51. The proteins in the eluate were detected by silver staining.

(L) (M) Quantitative analysis of RAD51 binding to MMS22L. Proteins were mixed with the indicated ratio, MMS22L immobilized on NiNTA beads and the eluate analyzed by silver staining. Lane 6: the same amount of RAD51 was used as in lane 5, but no MMS22L (No M) was added to control for non-specific binding of RAD51 to the resin. The proteins in the eluates were quantified by comparing with known amounts of MMS22L or RAD51, respectively. Averages shown, $n = 2$; error bars, SEM.

(N) A representative DNA strand exchange experiment with RAD51 (120 nM), ssDNA, RPA (30 nM) and MMS22L-TONSL. MMS22L-TONSL is not capable to alleviate the inhibitory effect of RPA on DNA strand exchange.

Figure S6. MMS22L-TONSL functions synergistically with RAD51 paralogues to promote RAD51 foci formation upon replication stress.

(A) Representative microscopy images showing the effect of MMS22L and BRCA2 depletion on the number of RAD51 foci formed upon CPT treatment (50 nM, 18 hours) in U2OS cells. Scale bar: 20 μm.

(B) Quantification of RAD51 to γH2AX foci ratios per nucleus ($n_{nuclei} \geq 50$) from a representative experiment in (A), $n = 3$.

(C, D, G) Quantification of RAD51 to γH2AX foci ratios per nucleus upon CPT treatment after depletion of MMS22L or TONSL with or without co-depletion of RAD51B, RAD51C and XRCC3. Similar effects are observed using various siRNA

oligonucleotides against MMS22L (C; $n = 2$) or TONSL (D; $n = 2$) in U2OS or HeLa cells (G; $n = 2$). Nuclei with ≥ 25 γH2AX foci (marker of replication stress sites) were analyzed.

(B-D) (G) Graphs represent distributions per nucleus. Boxes indicate the 25–75 percentile and whiskers the 10–90 percentile. Horizontal lines mark the medians. Statistical analysis: Mann–Whitney *U* test; *** $P \leq 0.0001$.

(E) (F) U2OS cells treated with the indicated siRNAs were collected and analyzed for MMS22L and RAD51 paralogues by western blotting (E) or qRT-PCR (F), $n = 3$; error bars, SEM. Note that depletion of RAD51C leads to decreased levels of XRCC3, as previously reported (Lio et al., 2004).

Experimental procedures

Cell culture conditions and reagents. U2OS and HeLa cell lines were grown in a humidified incubator at 37 °C and 5% CO_2 in Dulbecco's modified Eagle medium (DMEM, Gibco, Invitrogen) supplemented with 10% fetal bovine serum (FBS, PAA), 0.2 mM L-glutamine and standard antibiotics. The FRT-TetR-HeLa cell line (kindly provided by S. Taylor, University of Manchester) for creating stable cell lines using the Flp-In system (Invitrogen) was maintained as described previously(Tighe et al., 2008). For stable cell line generation, DNA plasmid transfections were performed using Lipofectamine 2000 (Invitrogen) according to the manufacturer instructions. The stable HeLa cell line for doxycycline-inducible downregulation of MMS22L was obtained by transfection of FRT-TetR-HeLa cells with the shMMS22L-expressing vector based on pSUPERIOR.puro (OligoEngine) and maintained in the presence of 0.5 μg/ml puromycin (Gibco). The stable cell lines for doxycycline-inducible expression of HSS-tagged TONSL constructs were obtained by transfecting FRT-TetR-HeLa cells with the pcDNA5 (Invitrogen)-based constructs of the respective TONSL variants and maintained in the presence of 100 μg/ml hygromycin B (Invitrogen). To induce expression in the FRT-TetR-HeLa system, 1 μg/ml doxycycline (Sigma-Aldrich) was used. Thymidine block and release experiments were performed as described previously(Piwko et al., 2010). In brief, cells were synchronized at the G1/S transition by 16 h incubation in medium containing 2 mM thymidine (Sigma-Aldrich), washed twice with pre-warmed DMEM, followed by incubation for 9 h. Subsequently, thymidine was added and cells were incubated for another 16 h. Next, cells were washed twice with DMEM and incubated 2 h to obtain S-phase cells and 11 h to obtain M (including 5 h of treatment with 0.23 μM nocodazole) and G1-synchronised cells. Subsequently, cells were treated with 5 μM etoposide for 3 h prior to fixation.
Cells were treated with the following drugs diluted in growth medium: Topo II inhibitor etoposide (5 μM; Sigma-Aldrich), Topo I inhibitor camptothecin (Sigma-Aldrich), hydroxyurea (Sigma-Aldrich), polymerase alpha inhibitor aphidicolin (2 μM; Sigma-Aldrich) and ATR inhibitor VE-821 (3 μM; Selleck Chemicals).

Cloning and DNA manipulations. The pCDNA5-HSS-TONSL plasmid for stable cell line generation by the Flip-In system (Invitrogen) was generated previously (Piwko et al., 2010). ΔLRR truncation was generated by PCR using primers: 5'-GACTGGATCCAATGAGCCTGGAGCGCGAGCTTC-3' and 5'-AGTCCTCGAGTCACAGGTCCCACGAAGTCACC-3' and sub-cloned into pcDNA5-HSS vector. LRR* mutation (double point mutation E1089K D1104N in the LRR domain) was obtained by QuikChange site-directed mutagenesis method (Stratagene) by two PCR reactions using oligonucleotide pairs: 5'-GACAAGTGTGTGGCTAAGCTGGTGGCTGCC-3' and 5'-GGCAGCCACCAGCTTAGCCACACACTTGTC-3', and 5'-GCCTGGCCCTCCTTAACCTCTCCTCCAATC-3' and 5'-GATTGGAGGAGAGGTTAAGGAGGGCCAGGC-3'. The plasmid for expressing shMMS22L was obtained from synthesized (Microsynth, Switzerland) annealed DNA fragments: 5'-GATCCCGGTTAGGTCTTCTGAATCATTCAAGAGATGATTCAGAAGACCTAACCTTTTTA-3' and 5'-AGCTTAAAAAGGTTAGGTCTTCTGAATCATCTCTTGAATGATTCAGAAGACCTAACCGG-3' cloned into pSUPERIOR.puro linearized with BglII and HindIII. The oligonucleotides corresponding to the target sequence of MMS22L are 5'-GGTTAGGTCTTCTGAATCA-3'. All obtained constructs were verified by DNA sequencing (Microsynth, Switzerland).

siRNA transfections. For RNAi experiments, HeLa or U2OS cells were transfected with 30 nM siRNAs using Lipofectamine RNAiMAX (Invitrogen) and incubated for 3 days prior analysis. All siRNA duplexes were purchased from Microsynth (Switzerland) with the exception of AllStars Negative Control siRNA, which was obtained from Qiagen. The target sequences of siRNA oligonucleotides used in this study were as follows: MMS22L (5'-GGUAGAAGAUGUUGCAAGU-3') (Piwko et al., 2010), MMS22L-4 (5'-CCCUUAAUGAUACGACGAA-3') (Piwko et al., 2010), TONSL (5'-GAGCUUGGCUAUUGU GGAU-3') (Piwko et al., 2010), CtIP (5'-GCUAAAACAGGAACGAAUC-3') SARTORI, 2007 #28}, RAD51 (5'-CAGGUGGUAGCUCAAGUGGAU-3'), RAD51B (5'-ACGAGUGGGUUUAUCACAAGA-

3')(Rodrigue et al., 2013), RAD51C (5'-CACCUUCUGUUCAGCACUAGA-3') (Rodrigue et al., 2013), XRCC3 (5'-CAGAAUUAUUGCUGCAAUUAA-3') (Rodrigue et al., 2013), BRCA2 (5'-GAAGCUGAUUCUCUGUCAU-3'), TONSL-3'UTR (pooled oligonucleotides 5'-CGACAGACCGAGACUCCGU-3' and 5'-CCUAAUAAAUGAAGCUGCU-3'). Downregulation of the targets was confirmed by immunoblotting and immunofluorescence using specific antibodies or quantitative RT-PCR.

Immunofluorescence microscopy. Cultured cells grown on cover slips were pre-extracted with 0.2% Triton-X100 in phosphate-buffered saline (PBS) for 3 min at room temperature before fixation with 4% formaldehyde in PBS for 12 min at room temperature, except for cells in Figure S1A, which were directly fixed with formaldehyde and subsequently permeabilized by incubation with 0.2% Triton-X100 for 5 minutes. Permeabilized cells were then incubated for 1 h at room temperature with blocking buffer (1.5% bovine serum albumin, 0.01% Triton X-100 in PBS), followed by additional 1 h incubation with antibodies diluted in blocking buffer. Secondary antibodies were labeled with Alexa Fluor 488 or 568 (Invitrogen). 0.2 mg/ml DAPI was added to visualize DNA. Samples were mounted using Immu-Mount (Thermo) and images were acquired on fully automated inverted epifluorescence microscopes (Ti-Eclipse, Nikon) with 60x or 100x oil objectives. A motorized XY-stage and piezo drive was used to acquire z-stacks (26 steps at 0.235 µm) and multiple fields of view per sample. Images in Figure 1A and Figure S1D were deconvolved using the Huygens deconvolution software (SVI). Endogenous MMS22L was stained using previously described antibody (Piwko et al., 2010). The following commercial antibodies were used for immunofluorescence: anti-γH2AX (05-636, Millipore, 1:750), anti-γH2AX (NB100-384, Novus Biologicals, 1:1,500), anti-RPA2 (9H8, Imgenex, 1:500), anti-BRCA2 (Clone 2B, Millipore, 1:300), anti-RAD51 (H-92, Santa Cruz Biotech, 1:400) and anti-RAD51 (14B4, Abcam, 1:500).

Image analysis. Non-deconvolved z-stack images were subjected to a maximum intensity projection using ImageJ software. CellProfiler (Carpenter et al., 2006) was used to analyze microscopy images (see Figure S1B). After

image illumination correction, global image thresholding using Otsu's method was applied to the DAPI channel to segment nuclei. Subsequently, nuclear foci in red and green channels were detected using Otsu's method and their signal intensities were quantified from illumination corrected images.

EM analysis. The procedure was essentially performed as previously described (Zellweger et al., 2015). Asynchronous subconfluent cultures of U2OS cells were treated with 25 nM CPT for 1 h. *In vivo* psoralen cross-linking of the DNA was achieved by a repetitive exposure of living cells to 4,5′,8-trimethylpsoralen (10 µg/ml final concentration) followed by irradiation pulses with UV 365-nm monochromatic light (UV Stratalinker 1800; Agilent Technologies). The cells were then lysed with cell lysis buffer (buffer C1: 1.28 M sucrose, 40 mM Tris-HCl [pH 7.5], 20 mM $MgCl_2$, and 4% Triton X-100; QIAGEN) and then digested by digestion buffer (QIAGEN buffer G2: 800 mM guanidine–HCl, 30 mM Tris–HCl [pH 8.0], 30 mM EDTA [pH 8.0], 5% Tween 20, and 0.5% Triton X-100) and 1 mg/ml proteinase K at 50 °C for 2 h. Chloroform/Isoamyl alcohol (24:1) was used to collect DNA via phase separation (centrifugation at 8,000 rpm for 20 min) followed by DNA precipitation by adding 0.7 x volume of isopropanol. The DNA was then washed with 70% ethanol, air dried, and resuspended in 200 µl TE (Tris-EDTA) buffer. 100 U restriction enzyme PvuII high-fidelity was used for 12 µg mammalian genomic DNA digestion (4–5-h incubation). Poly-Prep chromatography columns were used for RI enrichment. Benzoylated naphthoylated DEAE–cellulose granules were resuspended in 10 mM Tris-HCl (pH 8.0), and 300 mM NaCl to a final concentration of 0.1 g/ml. The columns were washed and equilibrated with 10 mM Tris-HCl (pH 8.0), and 1 M NaCl and 10 mM Tris-HCl (pH 8.0), and 300 mM NaCl, respectively. The sample DNA was then loaded and incubated for 0.5 h. After washing the columns (10 mM Tris-HCl [pH 8.0], and 1 M NaCl), the DNA was eluted in caffeine solution (10 mM Tris-HCl [pH 8.0], 1 M NaCl, and 1.8% [wt/vol] caffeine) for 10 min followed by sample collection. DNA is then purified and concentrated, using an Amicon size-exclusion column and resuspended in TE. With DNA spreading by the "BAC method", the DNA was loaded on carbon-coated 400-mesh copper grids. The DNA was then coated with platinum by platinum-carbon rotary shadowing (High

Vacuum Evaporator MED 020; Bal-Tec). Microscopy was performed with a transmission electron microscope (Tecnai G2 Spirit; FEI; LaB6 filament; high tension ≤ 120 kV) and acquired with a side mount charge-coupled device camera (2,600 × 4,000 pixels; Orius 1000; Gatan, Inc.). At last 70 replication fork molecules were analyzed for each experimental condition. The images were processed with DigitalMicrograph Version 1.83.842 (Gatan, Inc.) and analyzed with ImageJ (National Institutes of Health).

Immunoblotting, pulldowns and antibodies. For immunoblotting, proteins were resolved by SDS- PAGE and transferred to PVDF immobilon-P membrane (Millipore). Blocking and antibody incubations were carried out in 5% low-fat milk (Migros, Switzerland) in PBS-T (PBS, 0.1% Tween-20), and washings in PBS-T. Blots were developed using Immun-Star™ HRP (Biorad) and quantified using ImageJ software.

For HSS (2 x Strep-HA)-tagged-TONSL pulldown experiments from HeLa cell extracts, expression of indicated genes was induced by addition of doxycycline to medium for 24 h. Cell extracts were prepared as described previously (Piwko et al., 2010) except that lysis buffer was supplemented with 20 µg/ml avidin (IBA). Streptavidin MyOne T1 magnetic beads (Life Technologies) were used to capture Strep-tagged TONSL. The beads were washed four times with lysis buffer, bound proteins eluted by incubating with 2.5 mM D-biotin (Sigma-Aldrich), boiled at 95 °C and analyzed by immunoblotting.

Rabbit polyclonal antibodies against the 6 x His-tagged amino-terminal fragment (1-325) of MMS22L were described previously (Piwko et al., 2010). Rabbit polyclonal antibodies against TONSL were raised (Animal Facility, University of Zurich, Switzerland) and affinity-purified against a 6 x His-tagged fusion protein encoding amino acids 934-1012 of TONSL. The following commercial antibodies were used: anti-RAD51 (H-92, Santa Cruz Biotech, 1:2,000), anti-PCNA (PC10, Santa Cruz Biotech, 1:5,000), anti-RPA2 (9H8, Imgenex, 1:2,000), anti-MCM4 (clone B01P, Abnova, 1:1,000), anti-α-tubulin (DM1A, Sigma-Aldrich, 1:10,000), anti-HA.11 (16B12, Covance, 1:2,000), anti-ASF1A (C6E10, Cell Signaling, 1:5,000), anti-BRCA2 (Clone 2B, Millipore, 1:1,000), anti-GAPDH (MAB374, Millipore, 1:5,000), anti-XRCC3 (sc-53471, Santa Cruz Biotech, 1:1,000), anti-

RAD51C (SWE31 (Masson et al., 2001), a kind gift of J. Matos) and peroxidase-conjugated anti-mouse or anti-rabbit antibodies (Pierce, 1:5,000).

Quantitative real-time PCR. Total RNA was isolated from cells using the RNeasy Mini Kit (Qiagen). Reverse transcription was carried out using random primers and SuperScript II Reverse Transcriptase (Invitrogen) according to manufacturer's protocol. qRT-PCR was carried out using the LightCycler 480 (Roche Life Science) with the LightCycler 480 SYBR Green I Master (Roche) hot start reaction mix. Experiments were carried out in technical triplicates and normalised to GAPDH. The following primers were used: *CtIP* forward 5′-GTGAGGAAGACGTTATTCCAG-3′ and reverse 5′-TTTGCACACACAGAGTGCTC-3′; *RAD51B* forward 5'-CAAGAGCTGTGTGACCGTCTG-3' and reverse 5'-TCATGGACACCTCGATAACTCA-3' (Spandidos et al., 2008); GAPDH forward 5'-GAAGGTGAAGGTCGGAGTC-3' and reverse 5'-GAAGATGGTGATGGGATTTC-3'.

Flow cytometry cell cycle analysis. Synchronized HeLa cells were collected by trypsinization and fixed in 70% ethanol overnight at -20 °C. Cells were washed in PBS and incubated with 50 mg/ml propidium iodide and 20 mg/ml ribonuclease A for 30 min at room temperature. Flow cytometry was performed with a FACScalibur flow cytometer (BD Biosciences) using CellQuest software. Data analysis was performed using FlowJo software.

Preparation of recombinant proteins. The genes coding for MMS22L or TONSL were amplified from cDNA (Piwko et al., 2010) by PCR using primers hMMS22L-F (5′-GGCTAGCTGGGCCCGCTAGCGGATCCATGGAGAACTGTTCTGCT-3′) and hMMS22L-R (5′-CGCAAATCCTCGAGCCCGGGAGTATTATCATTTTCCAGTCTCT - 3′) for MMS22L and hTONSL2-F (5′-GGCTAGCTGGGCCCGCTAGCGGATCCATGAGCCTGGAGCGCGAGCTTC - 3′) and hTONSL-R (5′ - CGCAAATCCTCGAGCCCGGGTCAGAGGCGCCGAAAGAAGAGC - 3′) for TONSL. The PCR products of MMS22L were digested with ApaI and XhoI restriction endonucleases and cloned into ApaI and XhoI sites of pFB-MBP-Mlh3-his vector (Ranjha et al., 2014), generating pFB-MBP-MMS22L-his. The PCR

products of TONSL were digested with NheI and XhoI sites into pFB-GST-Top3 vector (Cejka and Kowalczykowski, 2010), generating pFB-GST-TONSL. The MMS22L-TONSL complex was expressed in *Sf*9 insect cells upon co-infection with MMS22L and TONSL baculoviruses according to manufacturer's recommendation (Invitrogen). The proteins were extracted from *Sf*9 cell pellets with 325 mM NaCl as described previously for Sgs1 (Cejka and Kowalczykowski, 2010). The protein was then bound to amylose resin (New England Biolabs) *via* the MBP tag on the MMS22L subunit. Both MBP and GST tags were then cleaved off by PreScission protease. The GST-tag was not used in our final purification protocol due to low binding efficiency of the recombinant construct to glutathione sepharose resin. The final protein complex was obtained by affinity purification on NiNTA agarose (Qiagen) *via* a his-tag on the C-terminus of MMS22L. The NaCl concentration was kept at 200 mM during the washing steps on both amylose and NiNTA resins to avoid disassembly of the complex. The MMS22L protein was prepared similarly, except 1M NaCl was used during the washing steps. The final sample was dialyzed against dialysis buffer containing 50 mM Tris-HCl (pH 7.5), 5 mM beta-mercaptoethanol, 100 mM NaCl, 10% glycerol, and 0.5 mM phenylmethanesulfonyl fluoride.

RAD51 was expressed from a pTXB3 vector (New England Biolabs) where it was fused to an intein tag at its C-terminus. Human RAD51 ORF was amplified by PCR from the plasmid pFB530 (Benson et al., 1994) using the following primers: 5'-GGGAATTCCATGGCAATGCAGATGCAGCTT-3' and 5'-AAGGCCGCTCTTCCGCAGTCTTTGGCATCTCCCACTCC-3'. The resulting PCR product was digested with NcoI and SapI and inserted in pTXB3 (New England Biolabs), creating a fusion of RAD51 with a self-cleavable affinity tag containing a chitin-binding domain. The fusion construct was expressed in *E. coli* BL21 cells and purified on Chitin binding resin (New England Biolabs) according to manufacturer's recommendation. The cleavage of the intein tag was carried out in a cleavage buffer containing 50 mM Tris-HCl (pH 8.0), 500 mM NaCl, 1 mM EDTA, 10% glycerol and 100 mM dithiothreitol for 32 h at 4 °C. The cleaved RAD51 was eluted with the same buffer without dithiothreitol, diluted with 50 mM Tris-HCl (pH 8.0) to a final concentration of 200 mM NaCl, and loaded onto pre-equilibrated HiTrap Q HP column (GE Healthcare). The column was washed

with R buffer containing 20 mM Tris-HCl (pH 8.0), 1 mM EDTA, 0.5 mM dithiothreitol, 10% glycerol supplemented with 150 mM NaCl. RAD51 was eluted with 10 ml gradient of 150 mM to 700 mM NaCl in R buffer supplemented with 1M NaCl. The final sample was dialyzed against dialysis buffer containing 20 mM Tris-HCl (pH 8.0), 20% glycerol, 100 mM NaCl, 1 mM dithiothreitol. Human RPA was purified as described previously (Thangavel et al., 2015).

Electrophoretic mobility shift assays. The binding reactions with RPA and MMS22L-TONSL (or MMS22L) were carried out in a 15 µl volume in 25 mM Tris-acetate (pH 7.5), 2 mM magnesium acetate, 1 mM dithiothreitol, 0.1 mg/ml bovine serum albumin (New England Biolabs), ^{32}P-labeled DNA substrate (1 nM, ssDNA, 5′ tailed DNA, or 3′ tailed DNA) and recombinant proteins as indicated for 30 min at 37 °C. Where indicated, RPA (8 nM for ssDNA, 64 nM for 3′ and 5′ tailed DNA) was pre-bound to the DNA on ice for 5 min; subsequently MMS22L-TONSL or MMS22L was added and the reaction was further incubated for 30 min at 37 °C. 5 µl of loading dye containing 50% glycerol and 0.01% bromophenol blue were added and the reactions were separated at 4 °C on 6% polyacrylamide gels in TBE buffer (89.1 mM Tris-Base, 88.6 mM boric acid, 2 mM EDTA) or on 4% polyacrylamide gels in TBE buffer for reaction containing RPA at 4 °C. The gels were dried, exposed to a Phosphorimager screen (GE Healthcare) and analyzed by Typhoon FLA 9500 (GE Healthcare). The ssDNA was oligonucleotide X12-3, the 3′ tailed DNA was prepared by annealing oligonucleotides X12-3 SC and X12-4 NC, the 5′ tailed DNA was prepared by annealing oligonucleotides X12-3 and X12-4 SC. The sequences of oligonucleotides were listed previously (Cejka and Kowalczykowski, 2010).

The binding assays (15 µl volume) with RAD51 and MMS22L-TONSL or MMS22L were performed similarly as described previously (Carreira et al., 2009). The binding buffer contained 20 mM Tris-HCl (pH 7.5), 10 mM magnesium acetate, 1 mM dithiothreitol, 0.1 mg/ml bovine serum albumin, 2 mM ATP, 2 mM calcium chloride, DNA substrate and recombinant proteins. For ssDNA reactions, we used unlabeled 3 nM poly dA_{40} oligonucleotide and 1 nM ^{32}P-radiolabeled poly dA_{40}. For dsDNA binding, the dA_{40} oligonucleotide was annealed with a complementary dT_{40}. We used 29 nM unlabeled poly dA_{40}:dT_{40}

and 1 nM ^{32}P-radiolabeled poly dA_{40}:dT_{40} dsDNA. The concentration of RAD51 was 150 nM in reactions with dsDNA and 640 nM in reactions with ssDNA. The reaction products were analyzed as described above.

Protein-protein pulldowns. Purified MMS22L-TONSL (with a MBP tag on the MMS22L subunit, 5.65 µg in a 150 µl reaction volume) was mixed with a 5 x excess of recombinant RAD51 or RPA in a binding buffer containing 50 mM Tris-HCl (pH 7.5), 2 mM beta-mercaptoethanol, 50 mM NaCl, 10% glycerol and 0.1% NP40. The reaction was mixed with 30 µl pre-equilibrated amylose resin (New England Biolabs) and incubated at room temperature for 30 min with occasional stirring. The resin was then washed four times with 300 µl binding buffer. Bound proteins were eluted with 60 µl protein loading buffer (50 mM Tris [pH 6.8], 10% glycerol, 1.6% sodium dodecyl sulfate, 0.1 M dithiothreitol, 0.01% bromophenol blue) by boiling at 95 °C for 3 min. The samples were analyzed on 10% SDS-PAGE and visualized by silver staining. Pulldowns with MMS22L (3.27 µg in a 210 µl reaction volume) were carried out as above, except for the binding buffer also contained 10 mM imidazole and NiNTA agarose (Qiagen) was used instead of amylose resin, exploiting the C-terminal his-tag on MMS22L. The resin was then washed four times with 450 µl binding buffer. Quantitative analysis of bound RAD51 was performed by a comparison with known amounts of the recombinant RAD51 or MMS22L-TONSL/MMS22L that were loaded on the same gel and used as protein standards, as described previously (Jensen et al., 2010).

Streptavidin or NiNTA pulldowns with biotinylated DNA, RPA and MMS22L-TONSL/MMS22L. 5′ biotinylated oligonucleotide (B-90-mer; 5′-CGGGTGTCGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGCGT-3′; 20 nM) was premixed with RPA (100 nM) in a 50 µl volume in a reaction buffer containing 50 mM Tris-HCl (pH 7.5), 120 mM NaCl, 0.1% (v/v) Triton X-100, 0.1 mg/ml bovine serum albumin and incubated for 15 min at 37 °C. Then MMS22L-TONSL or MMS22L (both 50 nM) were added and incubated for further 15 min at 37 °C. In reactions lacking RPA or MMS22L-TONSL/MMS22L, the reaction was mock incubated under the same experimental conditions. Reaction mixture was subsequently

mixed with 15 µl pre-equilibrated Streptavidin Dynabeads (Invitrogen, M-280) or NiNTA agarose (Qiagen) and incubated at 4 °C for 30 min. The mixture was stirred occasionally. The beads were washed four times with 150 µl wash buffer containing 50 mM Tris-HCl (pH 7.5), 120 mM NaCl, 0.1% (v/v) Triton X-100 and eluted with 60 µl protein loading buffer (50 mM Tris [pH 6.8], 10% glycerol, 1.6% sodium dodecyl sulfate, 0.1 M dithiothreitol, 0.01% bromophenol blue) by boiling for 3 min at 95 °C. Samples were resolved on 10 % SDS-polyacrylamide gel. The proteins in the pulldowns were detected by western blotting using anti-RPA antibody (NA19L; Calbiochem; 1:500) or anti-his tag antibody (A00186-100; GenScript; 1:2,500) against his-tagged MMS22L.

Streptavidin pulldowns with biotinylated ssDNA, RAD51 and MMS22L-TONSL/MMS22L. The assay was performed essentially as described previously (Jensen et al., 2010) in a 20 µl reaction volume. RAD51 (200 nM) alone or with MMS22L-TONSL or MMS22L (20, 80, and 150 nM) were pre-incubated at 37 °C for 15 min in a binding buffer containing 20 mM Tris-HCl (pH 7.5), 120 mM NaCl, 0.1% (v/v) Triton X-100, 2 mM calcium chloride, 10 mM magnesium acetate, 2 mM ATP, 1 mM dithiothreitol and 0.1 mg/ml bovine serum albumin. Then biotinylated B-90-mer (for sequence see above; 1 nM) was added with or without cold dsDNA (X12-3 annealed with X12-4 C (Cejka and Kowalczykowski, 2010); 10 nM) and incubated for 5 min at 37 °C. Reactions were further incubated upon adding pre-equilibrated Streptavidin Dynabeads (Invitrogen, M-280) at 4 °C for 30 min while stirring occasionally. The beads were washed four times with 150 µl wash buffer containing 20 mM Tris-HCl (pH 7.5), 120 mM NaCl, 0.1% (v/v) Triton X-100, 2 mM calcium chloride, 10 mM magnesium acetate, 2 mM ATP and 1 mM dithiothreitol. The protein was eluted by boiling with protein loading buffer (50 mM Tris [pH 6.8], 10% glycerol, 1.6% sodium dodecyl sulfate, 0.1 M dithiothreitol, 0.01% bromophenol blue) and 1 µl bovine serum albumin (10 mg/ml) at 95 °C for 3 min and loaded onto 10% SDS-polyacrylamide gel. The RAD51 protein bound to ssDNA was detected by western blotting using a polyclonal anti-RAD51 antibody (Simandlova et al., 2013).

Strand exchange assays. The reactions (15 µl) were performed essentially as described previously (Jensen et al., 2010). First, non-labeled DNA substrate (2.5 nM, either ssDNA, 5′ tailed, 3′ tailed or gapped DNA) was pre-incubated with the indicated amounts of RAD51 in a binding buffer containing 25 mM Tris-acetate (pH 7.5), 3 mM magnesium acetate, 2 mM calcium chloride, 0.1 mg/ml bovine serum albumin, 1 mM ATP and 1 mM DTT for 5 min at 37 °C. Subsequently, MMS22L-TONSL or MMS22L was added to the reaction mixture and incubated for 5 min at 37 °C. Next, ^{32}P-radiolabeled dsDNA (2.5 nM) was added and the reaction was further incubated for 30 min at 37 °C. The reaction was terminated by the addition of 5 µl STOP buffer (150 mM EDTA, 2% sodium dodecyl sulfate, 30% glycerol, 0.1% bromophenol blue) and 1 µl proteinase K (14-24 mg/ml, Roche) for 10 min at 37 °C. The reaction products were separated on 8% polyacrylamide gels in TBE buffer (89.1 mM Tris-Base, 88.6 mM boric acid, 2 mM EDTA). Where indicated, RPA was pre-bound to the DNA for 5 min at 37 °C in the binding buffer containing 25 mM Tris-acetate (pH 7.5), 0.66 mM magnesium acetate, 2 mM calcium chloride, 0.1 mg/ml bovine serum albumin, 1 mM ATP, and 1 mM DTT; subsequently MMS22L-TONSL or MMS22L was added and incubated for 5 min at 37 °C. Upon addition of ^{32}P-radiolabeled dsDNA (2.5 nM) the reaction was further incubated for 30 min at 37 °C. The gel was dried and exposed to a Phosphorimager screen (GE Healthcare) and analyzed by Typhoon FLA 9500 (GE Healthcare). The ssDNA was oligonucleotide RJ-167-mer, the 3′ tailed DNA was prepared by annealing oligonucleotides RJ-167-mer and RJ-PHIX-42-1, the 5′ tailed DNA was prepared by annealing oligonucleotide RJ-5′TAIL-167-mer and RJ-PHIX-42-1, the gapped DNA was prepared by annealing oligonucleotide RJ-167-mer, RJ-PHIX-42-1 and RJ-END, the dsDNA was prepared by annealing oligonucleotide RJ-Oligo1 and RJ-Oligo2. The sequences of oligonucleotides were described previously (Jensen et al., 2010), except RJ-END (5′-TAAATAAGATAAGGA-3′).

iPOND assay. For iPOND (Sirbu et al., 2011), HeLa cells were incubated in DMEM containing 10 µM EdU (Invitrogen) for 10 min. For the CPT-recovery assay, EdU-labeled cells were first treated with 1 µM CPT for 30 min, then washed three times with 10 mM thymidine-containing medium and incubated

for indicated times. For the experiment in Figures 3D and 3E, stably transfected HeLa cells with doxycycline-inducible shMMS22L were treated with doxycycline for 48 h before performing CPT-recovery iPOND. For the experiment in Figures S3B and S3C, endogenous TONSL was depleted by siRNA targeting 3'UTR for 24 h and TONSL variants were expressed from a Dox-inducible promoter for additional 24 h before the CPT-recovery assay. To cross-link proteins to DNA, cells were incubated in PBS containing 1% formaldehyde for 20 min at room temperature. To quench excess of cross-linker, glycine was added to a final concentration of 125 mM for 10 min. Cells were washed three times with PBS and collected by centrifugation. For cell permeabilization, pellets were resuspended in PBS containing 0.25% Triton X-100 and complete protease inhibitor cocktail (Roche), and incubated at 4 °C for 20 min. Permeabilized cells were washed once with PBS containing 0.5% bovine serum albumin (Roche). Cells were incubated in Click reaction buffer (PBS containing 10 mM sodium ascorbate (Sigma-Aldrich), 10 μM Azide-PEG11-Biotin Conjugate (Tocris Bioscience) and 2 μM $CuSO_4$) for 90 min at room temperature. Cells were washed twice with PBS containing 0.5% bovine serum albumin and resuspended in lysis buffer (25 mM Tris-HCl [pH 7.0], 150 mM NaCl, 0.2% NP-40, 1% sodium deoxycholate, 0.1% sodium dodecyl sulfate, complete protease inhibitor cocktail, Roche). Suspension was sonicated on ice using Branson 250 sonifier, and then centrifuged for 30 min at 20,800 x *g*. Supernatants were collected and Streptavidin MyOne T1 magnetic beads (Life Technologies) was added and incubated for 1 h at 4 °C. Beads were washed twice with lysis buffer, once with 1 M NaCl, and finally once with lysis buffer. Captured proteins were eluted by boiling for 30 min at 95 °C and detected by western blotting.

Statistical analysis. Statistical analyses were performed and standard error of the mean values were calculated using Prism software (GraphPad). Mann–Whitney *U* and Pearson correlation test were employed to verify statistical significance.

References

Benson, F.E., Stasiak, A., and West, S.C. (1994). Purification and characterization of the human Rad51 protein, an analogue of E. coli RecA. The EMBO journal *13*, 5764-5771.

Carpenter, A.E., Jones, T.R., Lamprecht, M.R., Clarke, C., Kang, I.H., Friman, O., Guertin, D.A., Chang, J.H., Lindquist, R.A., Moffat, J., *et al.* (2006). CellProfiler: image analysis software for identifying and quantifying cell phenotypes. Genome biology *7*, R100.

Carreira, A., Hilario, J., Amitani, I., Baskin, R.J., Shivji, M.K., Venkitaraman, A.R., and Kowalczykowski, S.C. (2009). The BRC repeats of BRCA2 modulate the DNA-binding selectivity of RAD51. Cell *136*, 1032-1043.

Cejka, P., and Kowalczykowski, S.C. (2010). The full-length Saccharomyces cerevisiae Sgs1 protein is a vigorous DNA helicase that preferentially unwinds holliday junctions. The Journal of biological chemistry *285*, 8290-8301.

Cimprich, K.A., and Cortez, D. (2008). ATR: an essential regulator of genome integrity. Nature reviews. Molecular cell biology *9*, 616-627.

Collins, S.R., Miller, K.M., Maas, N.L., Roguev, A., Fillingham, J., Chu, C.S., Schuldiner, M., Gebbia, M., Recht, J., Shales, M., *et al.* (2007). Functional dissection of protein complexes involved in yeast chromosome biology using a genetic interaction map. Nature *446*, 806-810.

Duro, E., Lundin, C., Ask, K., Sanchez-Pulido, L., MacArtney, T.J., Toth, R., Ponting, C.P., Groth, A., Helleday, T., and Rouse, J. (2010). Identification of the MMS22L-TONSL complex that promotes homologous recombination. Molecular cell *40*, 632-644.

Duro, E., Vaisica, J.A., Brown, G.W., and Rouse, J. (2008). Budding yeast Mms22 and Mms1 regulate homologous recombination induced by replisome blockage. DNA repair *7*, 811-818.

Forbes, S.A., Beare, D., Gunasekaran, P., Leung, K., Bindal, N., Boutselakis, H., Ding, M., Bamford, S., Cole, C., Ward, S., *et al.* (2015). COSMIC: exploring the world's knowledge of somatic mutations in human cancer. Nucleic acids research *43*, D805-811.

Golmard, L., Caux-Moncoutier, V., Davy, G., Al Ageeli, E., Poirot, B., Tirapo, C., Michaux, D., Barbaroux, C., d'Enghien, C.D., Nicolas, A., *et al.* (2013). Germline mutation in the RAD51B gene confers predisposition to breast cancer. BMC cancer *13*, 484.

Jensen, R.B., Carreira, A., and Kowalczykowski, S.C. (2010). Purified human BRCA2 stimulates RAD51-mediated recombination. Nature *467*, 678-683.

Jensen, R.B., Ozes, A., Kim, T., Estep, A., and Kowalczykowski, S.C. (2013). BRCA2 is epistatic to the RAD51 paralogs in response to DNA damage. DNA repair *12*, 306-311.

Lio, Y.C., Schild, D., Brenneman, M.A., Redpath, J.L., and Chen, D.J. (2004). Human Rad51C deficiency destabilizes XRCC3, impairs recombination, and radiosensitizes S/G2-phase cells. The Journal of biological chemistry *279*, 42313-42320.

Liu, J., Doty, T., Gibson, B., and Heyer, W.D. (2010). Human BRCA2 protein promotes RAD51 filament formation on RPA-covered single-stranded DNA. Nature structural & molecular biology *17*, 1260-1262.

Liu, J., Renault, L., Veaute, X., Fabre, F., Stahlberg, H., and Heyer, W.D. (2011). Rad51 paralogues Rad55-Rad57 balance the antirecombinase Srs2 in Rad51 filament formation. Nature *479*, 245-248.

Luke, B., Versini, G., Jaquenoud, M., Zaidi, I.W., Kurz, T., Pintard, L., Pasero, P., and Peter, M. (2006). The cullin Rtt101p promotes replication fork progression through damaged DNA and natural pause sites. Current biology : CB *16*, 786-792.

Masson, J.Y., Stasiak, A.Z., Stasiak, A., Benson, F.E., and West, S.C. (2001). Complex formation by the human RAD51C and XRCC3 recombination repair proteins. Proceedings of the National Academy of Sciences of the United States of America *98*, 8440-8446.

Meindl, A., Hellebrand, H., Wiek, C., Erven, V., Wappenschmidt, B., Niederacher, D., Freund, M., Lichtner, P., Hartmann, L., Schaal, H., *et al.* (2010). Germline mutations in breast and ovarian cancer pedigrees establish RAD51C as a human cancer susceptibility gene. Nature genetics *42*, 410-414.

Neelsen, K.J., and Lopes, M. (2015). Replication fork reversal in eukaryotes: from dead end to dynamic response. Nature reviews. Molecular cell biology *16*, 207-220.

Nitiss, J.L. (2009). Targeting DNA topoisomerase II in cancer chemotherapy. Nature reviews. Cancer *9*, 338-350.

O'Connell, B.C., Adamson, B., Lydeard, J.R., Sowa, M.E., Ciccia, A., Bredemeyer, A.L., Schlabach, M., Gygi, S.P., Elledge, S.J., and Harper, J.W. (2010). A genome-wide camptothecin sensitivity screen identifies a mammalian MMS22L-NFKBIL2 complex required for genomic stability. Molecular cell *40*, 645-657.

O'Donnell, L., Panier, S., Wildenhain, J., Tkach, J.M., Al-Hakim, A., Landry, M.C., Escribano-Diaz, C., Szilard, R.K., Young, J.T., Munro, M., *et al.* (2010). The MMS22L-TONSL complex mediates recovery from replication stress and homologous recombination. Molecular cell *40*, 619-631.

Petermann, E., and Helleday, T. (2010). Pathways of mammalian replication fork restart. Nature reviews. Molecular cell biology *11*, 683-687.

Petermann, E., Orta, M.L., Issaeva, N., Schultz, N., and Helleday, T. (2010). Hydroxyurea-stalled replication forks become progressively inactivated and require two different RAD51-mediated pathways for restart and repair. Molecular cell *37*, 492-502.

Piwko, W., Olma, M.H., Held, M., Bianco, J.N., Pedrioli, P.G., Hofmann, K., Pasero, P., Gerlich, D.W., and Peter, M. (2010). RNAi-based screening identifies the Mms22L-Nfkbil2 complex as a novel regulator of DNA replication in human cells. The EMBO journal *29*, 4210-4222.

Qing, Y., Yamazoe, M., Hirota, K., Dejsuphong, D., Sakai, W., Yamamoto, K.N., Bishop, D.K., Wu, X., and Takeda, S. (2011). The epistatic relationship between BRCA2 and the other RAD51 mediators in homologous recombination. PLoS genetics *7*, e1002148.

Raderschall, E., Golub, E.I., and Haaf, T. (1999). Nuclear foci of mammalian recombination proteins are located at single-stranded DNA regions formed after DNA damage. Proceedings of the National Academy of Sciences of the United States of America *96*, 1921-1926.

Ranjha, L., Anand, R., and Cejka, P. (2014). The Saccharomyces cerevisiae Mlh1-Mlh3 heterodimer is an endonuclease that preferentially binds to Holliday junctions. The Journal of biological chemistry *289*, 5674-5686.

Ray Chaudhuri, A., Hashimoto, Y., Herrador, R., Neelsen, K.J., Fachinetti, D., Bermejo, R., Cocito, A., Costanzo, V., and Lopes, M. (2012). Topoisomerase I poisoning results in PARP-mediated replication fork reversal. Nature structural & molecular biology *19*, 417-423.

Rodrigue, A., Coulombe, Y., Jacquet, K., Gagne, J.P., Roques, C., Gobeil, S., Poirier, G., and Masson, J.Y. (2013). The RAD51 paralogs ensure cellular protection against mitotic defects and aneuploidy. Journal of cell science *126*, 348-359.

San Filippo, J., Sung, P., and Klein, H. (2008). Mechanism of eukaryotic homologous recombination. Annual review of biochemistry *77*, 229-257.

Sartori, A.A., Lukas, C., Coates, J., Mistrik, M., Fu, S., Bartek, J., Baer, R., Lukas, J., and Jackson, S.P. (2007). Human CtIP promotes DNA end resection. Nature *450*, 509-514.

Sigurdsson, S., Trujillo, K., Song, B., Stratton, S., and Sung, P. (2001a). Basis for avid homologous DNA strand exchange by human Rad51 and RPA. The Journal of biological chemistry *276*, 8798-8806.

Sigurdsson, S., Van Komen, S., Bussen, W., Schild, D., Albala, J.S., and Sung, P. (2001b). Mediator function of the human Rad51B-Rad51C complex in Rad51/RPA-catalyzed DNA strand exchange. Genes & development *15*, 3308-3318.

Simandlova, J., Zagelbaum, J., Payne, M.J., Chu, W.K., Shevelev, I., Hanada, K., Chatterjee, S., Reid, D.A., Liu, Y., Janscak, P., *et al.* (2013). FBH1 helicase disrupts

RAD51 filaments in vitro and modulates homologous recombination in mammalian cells. The Journal of biological chemistry *288*, 34168-34180.

Sirbu, B.M., Couch, F.B., Feigerle, J.T., Bhaskara, S., Hiebert, S.W., and Cortez, D. (2011). Analysis of protein dynamics at active, stalled, and collapsed replication forks. Genes & development *25*, 1320-1327.

Spandidos, A., Wang, X., Wang, H., Dragnev, S., Thurber, T., and Seed, B. (2008). A comprehensive collection of experimentally validated primers for Polymerase Chain Reaction quantitation of murine transcript abundance. BMC genomics *9*, 633.

Sung, P. (1994). Catalysis of ATP-dependent homologous DNA pairing and strand exchange by yeast RAD51 protein. Science *265*, 1241-1243.

Tarsounas, M., Davies, A.A., and West, S.C. (2004). RAD51 localization and activation following DNA damage. Philosophical transactions of the Royal Society of London. Series B, Biological sciences *359*, 87-93.

Taylor, M.R., Spirek, M., Chaurasiya, K.R., Ward, J.D., Carzaniga, R., Yu, X., Egelman, E.H., Collinson, L.M., Rueda, D., Krejci, L., *et al.* (2015). Rad51 Paralogs Remodel Pre-synaptic Rad51 Filaments to Stimulate Homologous Recombination. Cell *162*, 271-286.

Thangavel, S., Berti, M., Levikova, M., Pinto, C., Gomathinayagam, S., Vujanovic, M., Zellweger, R., Moore, H., Lee, E.H., Hendrickson, E.A., *et al.* (2015). DNA2 drives processing and restart of reversed replication forks in human cells. The Journal of cell biology *208*, 545-562.

Thorslund, T., McIlwraith, M.J., Compton, S.A., Lekomtsev, S., Petronczki, M., Griffith, J.D., and West, S.C. (2010). The breast cancer tumor suppressor BRCA2 promotes the specific targeting of RAD51 to single-stranded DNA. Nature structural & molecular biology *17*, 1263-1265.

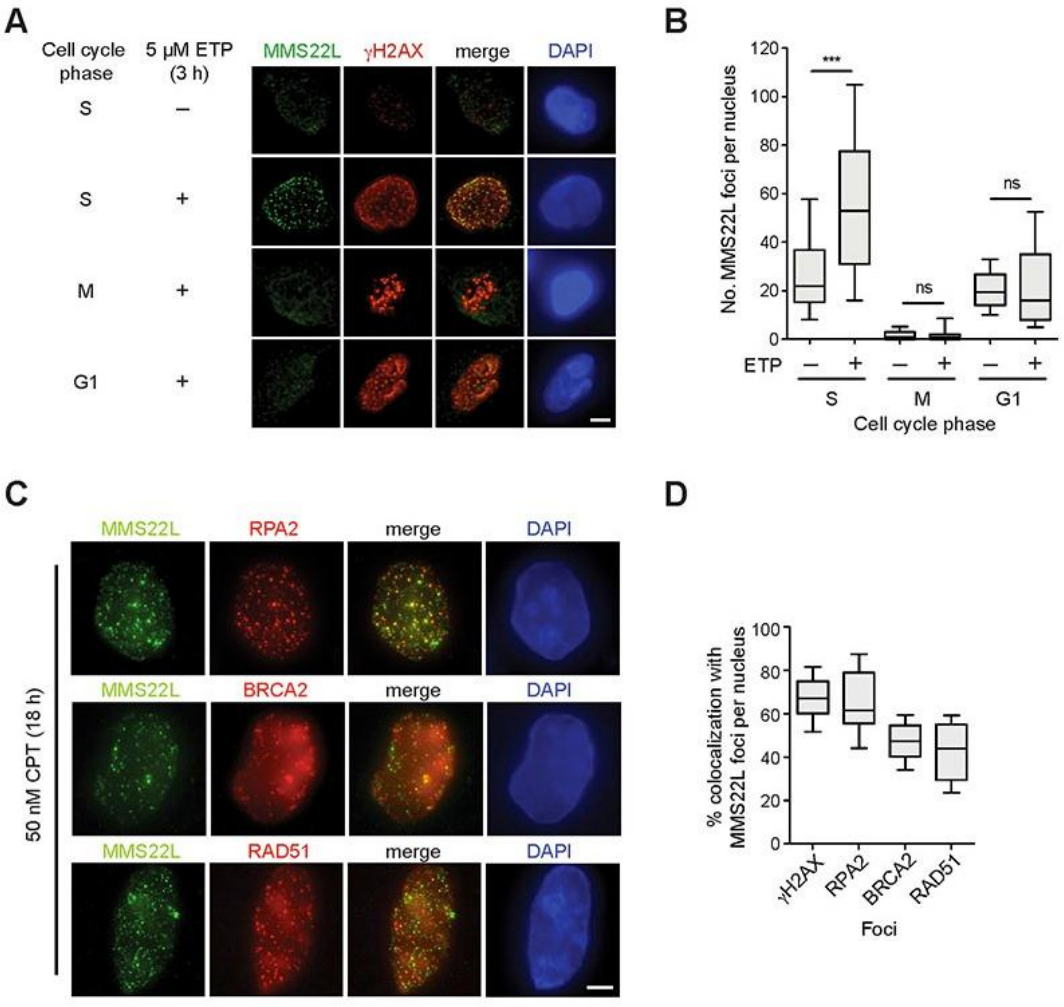
Tighe, A., Staples, O., and Taylor, S. (2008). Mps1 kinase activity restrains anaphase during an unperturbed mitosis and targets Mad2 to kinetochores. The Journal of cell biology *181*, 893-901.

Wooster, R., Bignell, G., Lancaster, J., Swift, S., Seal, S., Mangion, J., Collins, N., Gregory, S., Gumbs, C., and Micklem, G. (1995). Identification of the breast cancer susceptibility gene BRCA2. Nature *378*, 789-792.

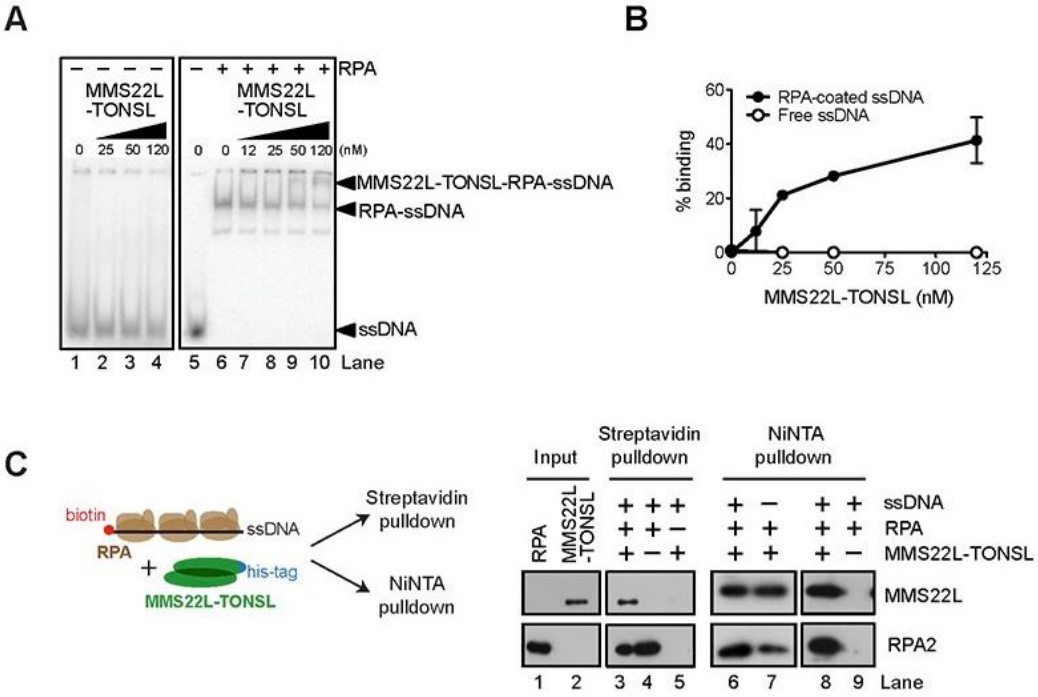
Zaidi, I.W., Rabut, G., Poveda, A., Scheel, H., Malmstrom, J., Ulrich, H., Hofmann, K., Pasero, P., Peter, M., and Luke, B. (2008). Rtt101 and Mms1 in budding yeast form a CUL4(DDB1)-like ubiquitin ligase that promotes replication through damaged DNA. EMBO reports *9*, 1034-1040.

Zellweger, R., Dalcher, D., Mutreja, K., Berti, M., Schmid, J.A., Herrador, R., Vindigni, A., and Lopes, M. (2015). Rad51-mediated replication fork reversal is a global response to genotoxic treatments in human cells. The Journal of cell biology *208*, 563-579.

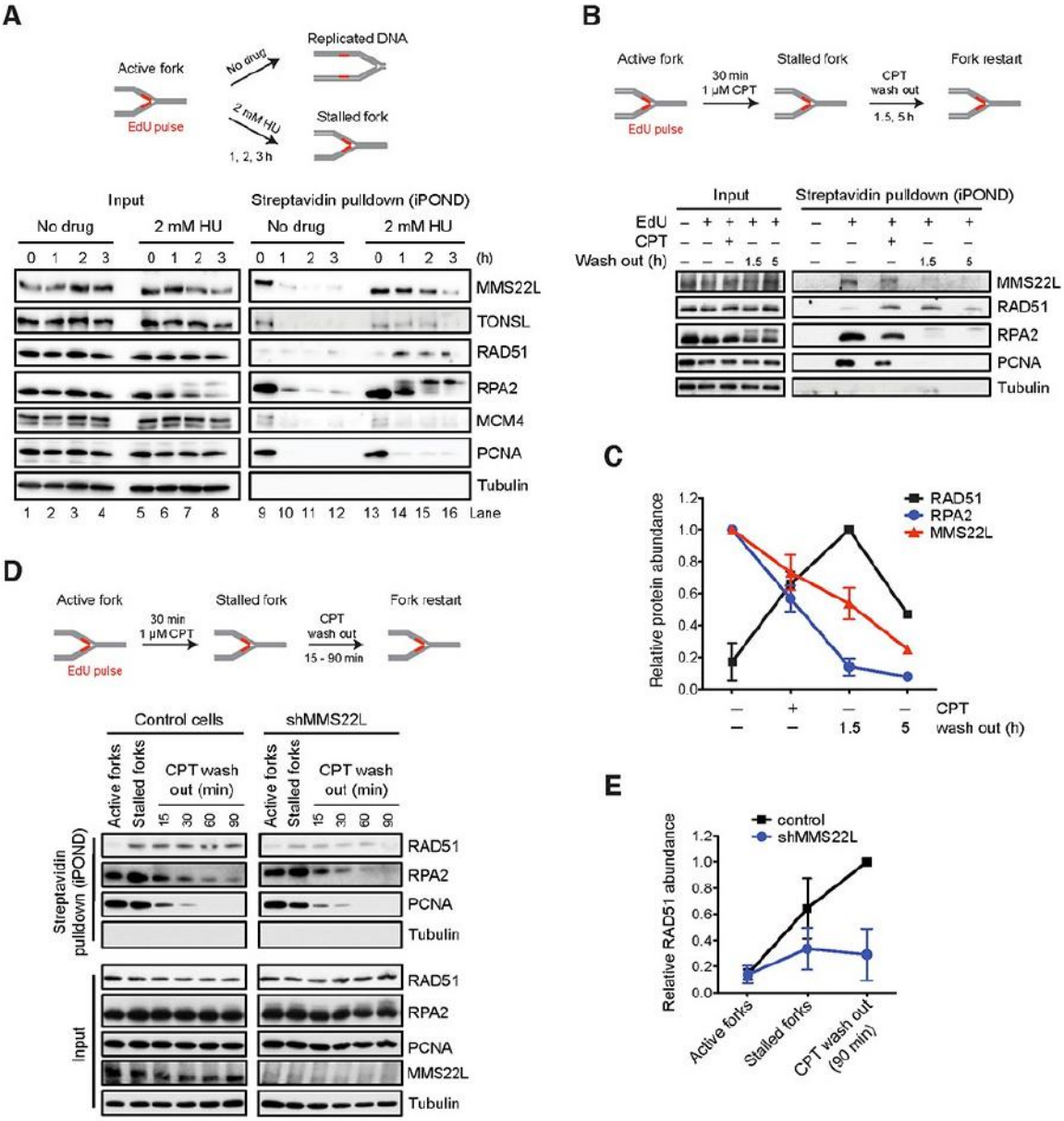
Piwko et al., Figure 1



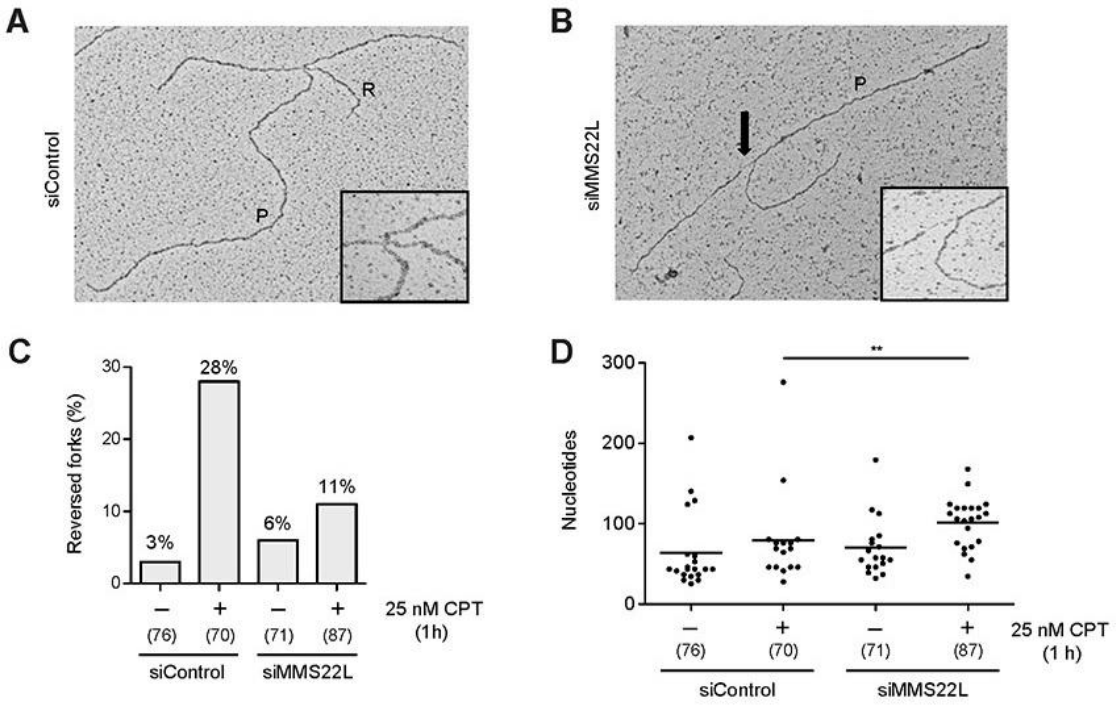
Piwko et al., Figure 2



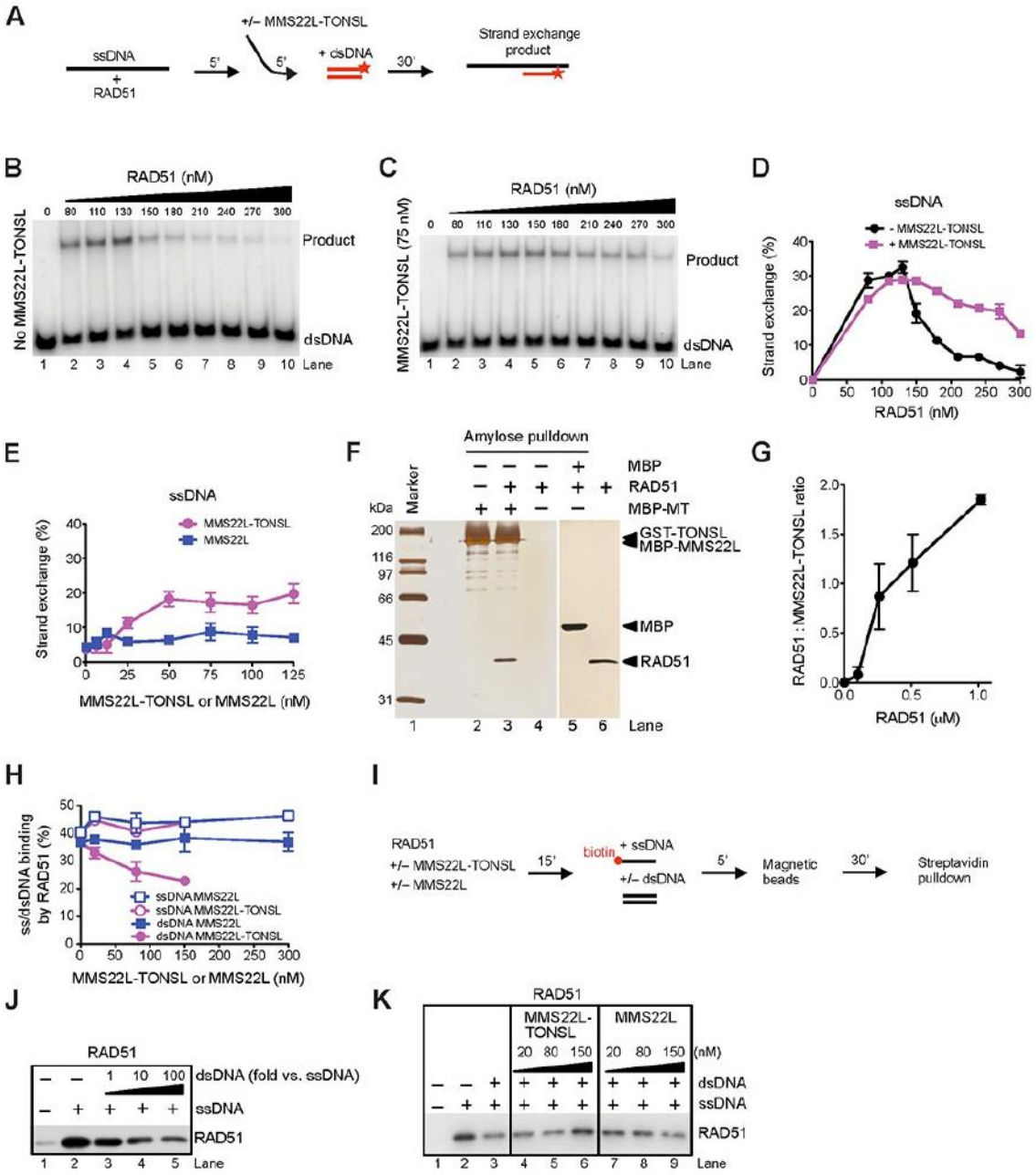
Piwko et al., Figure 3



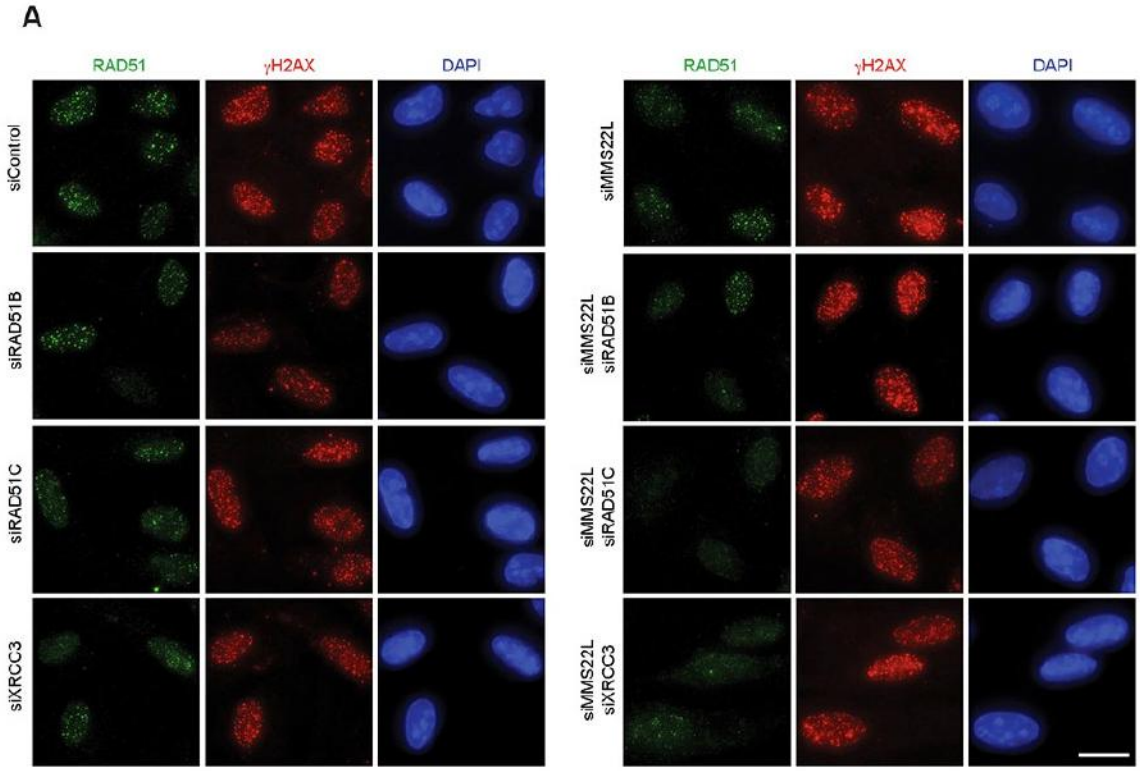
Piwko et al., Figure 4

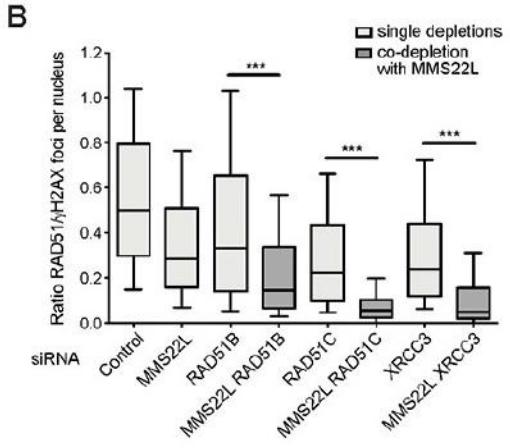


Piwko et al., Figure 5

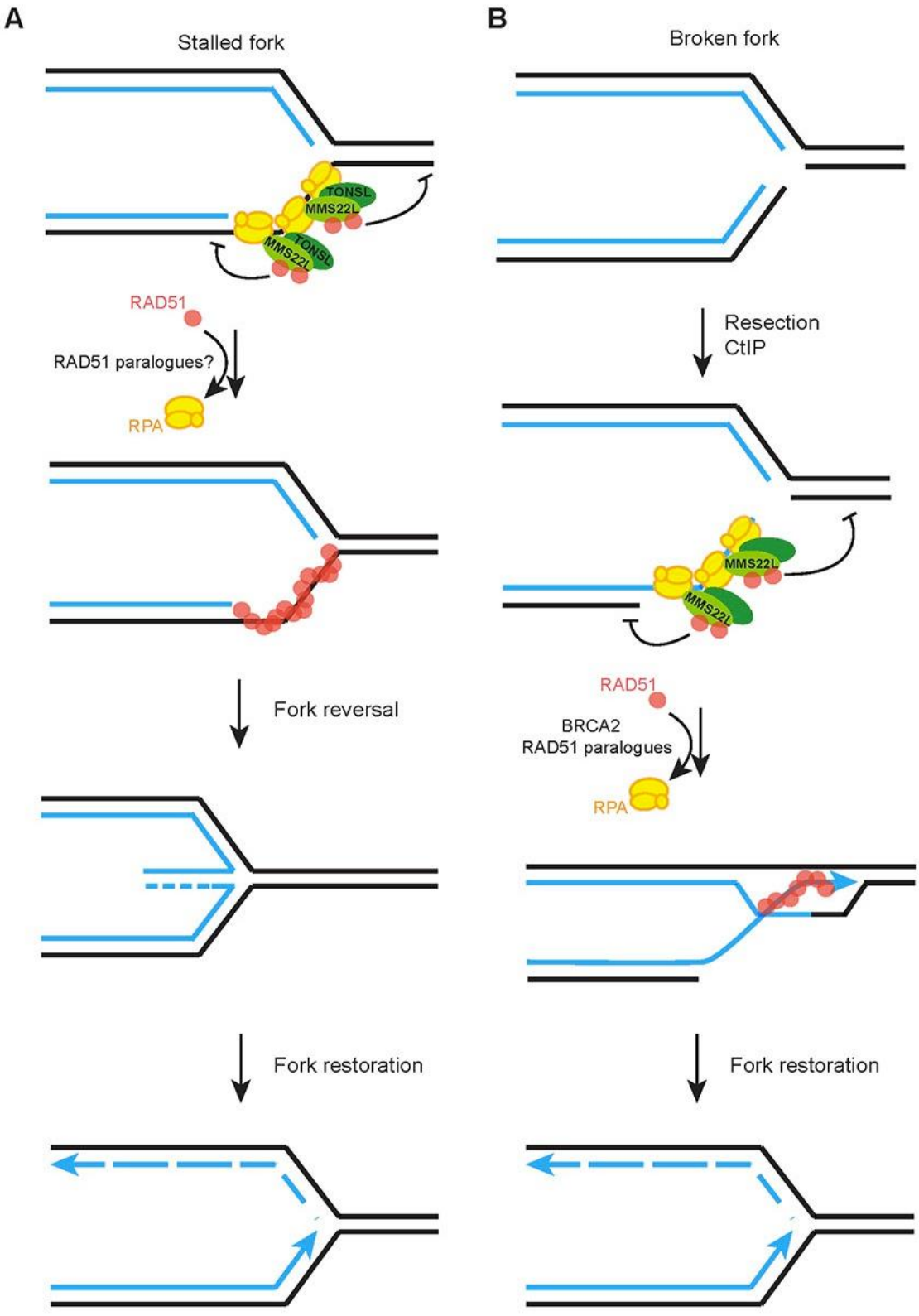


Piwko et al., Figure 6

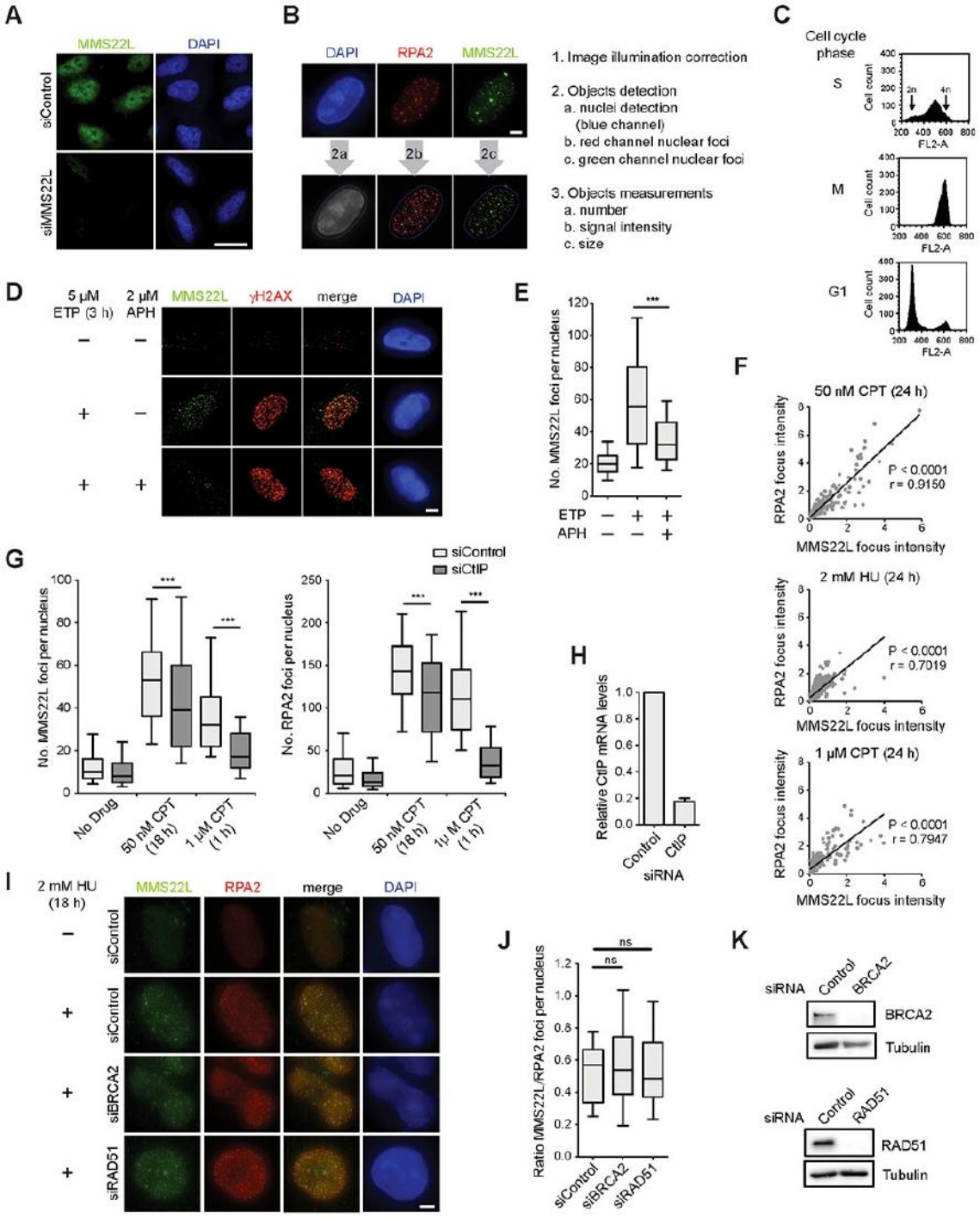




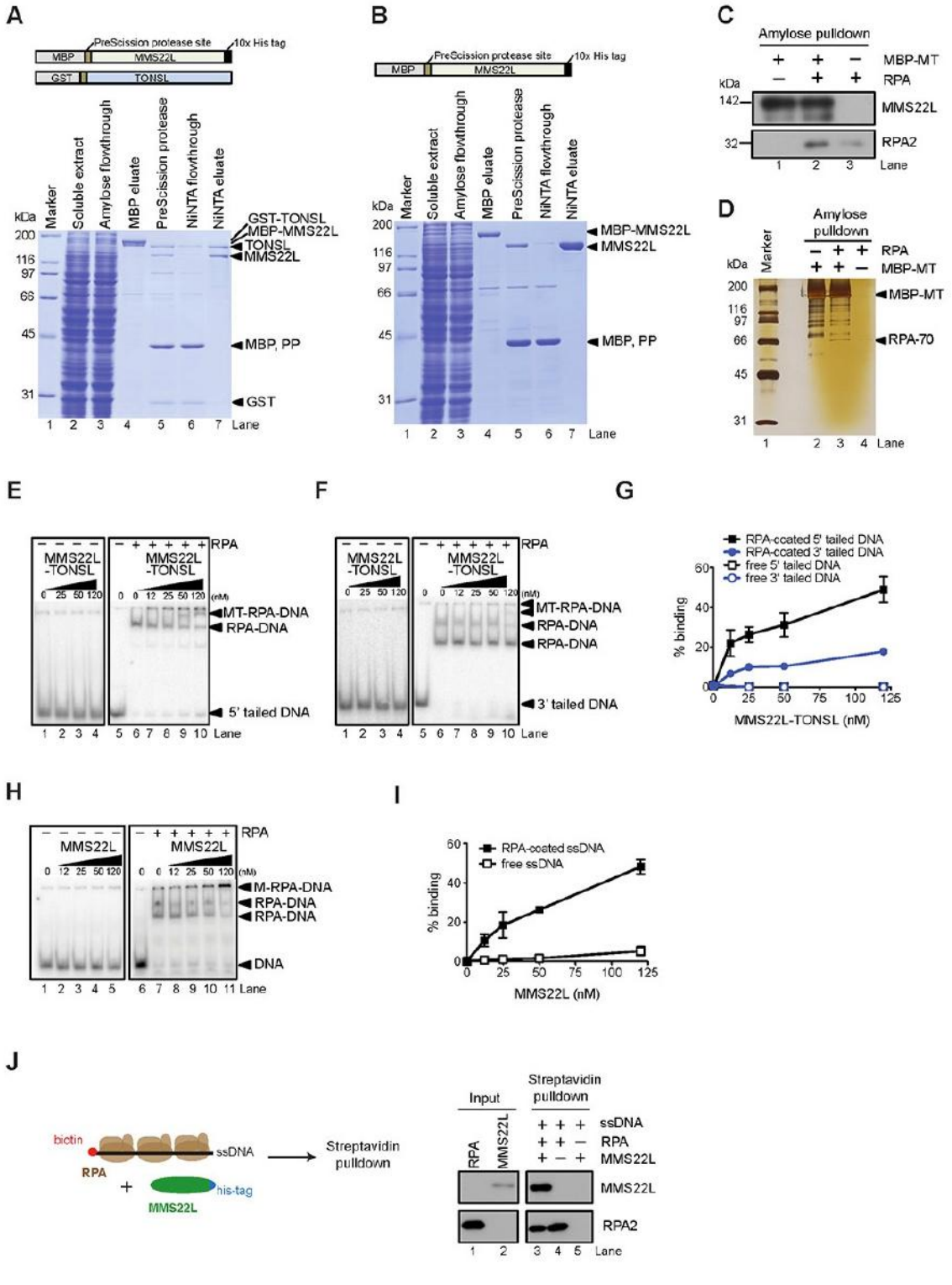
Piwko et al., Figure 7



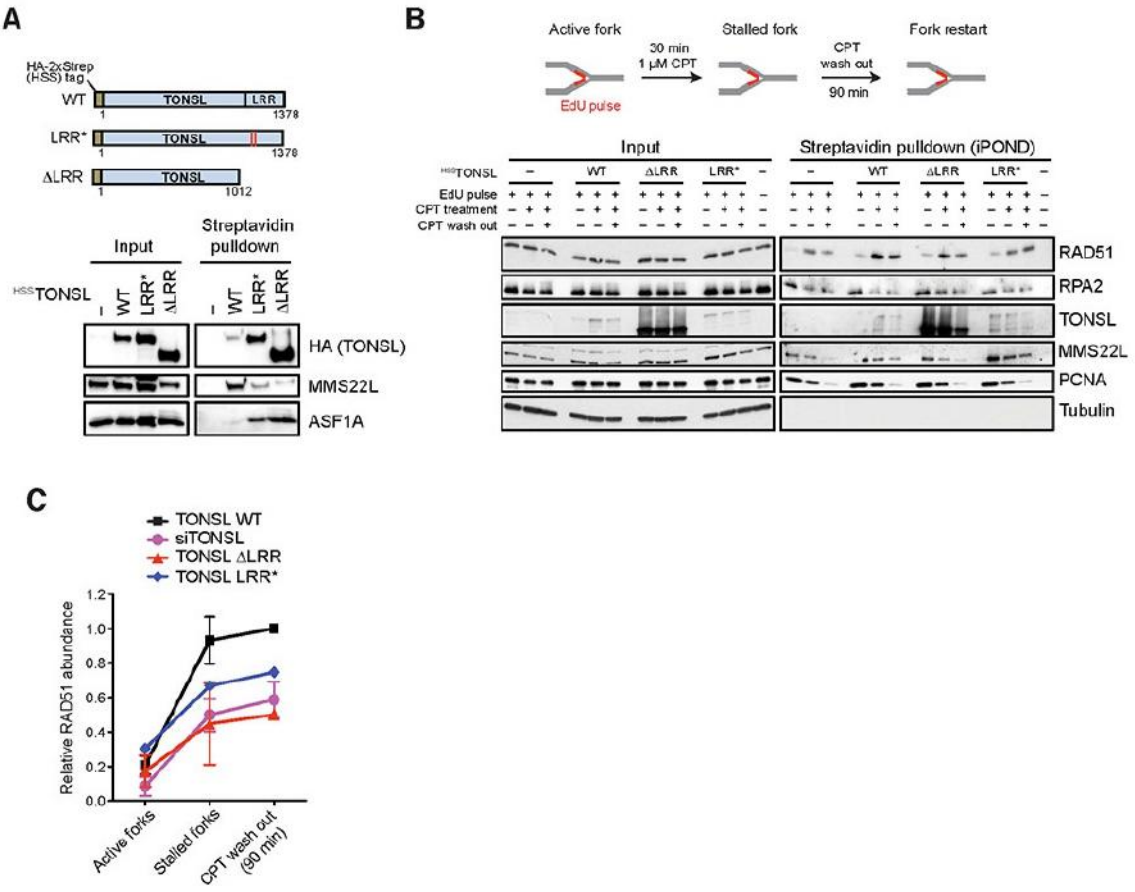
Piwko et al., Figure S1



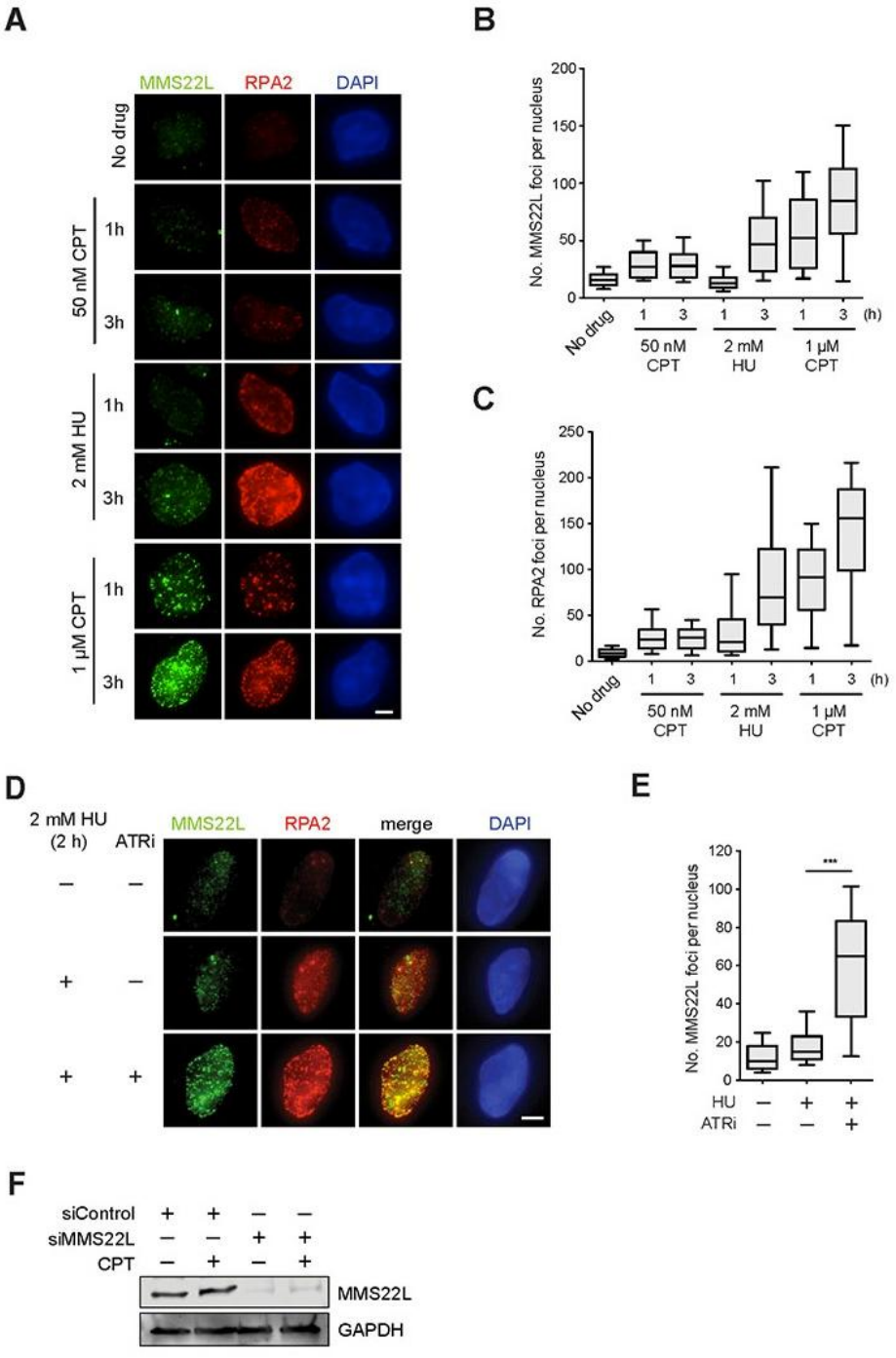
Piwko et al., Figure S2



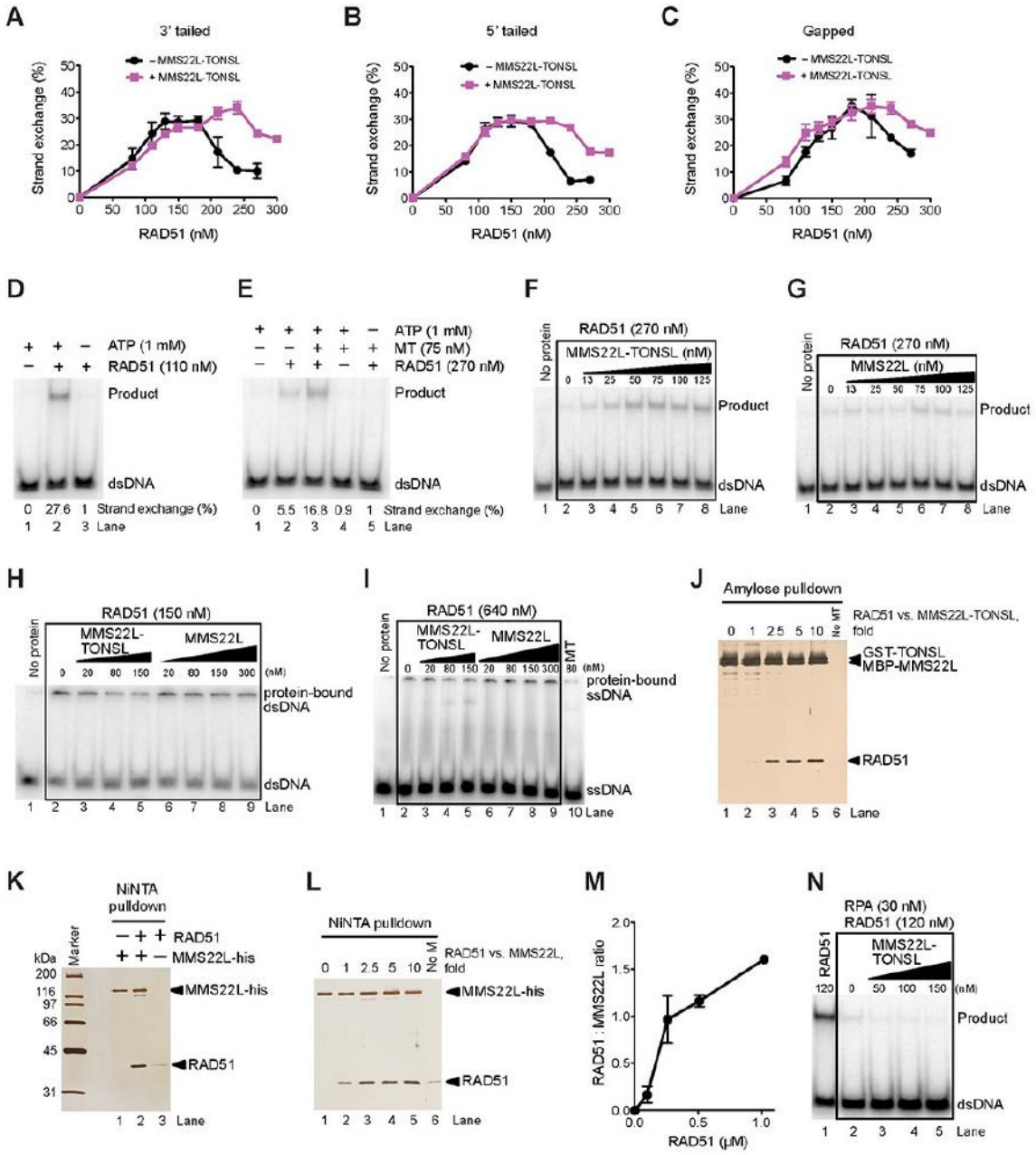
Piwko et al., Figure S3



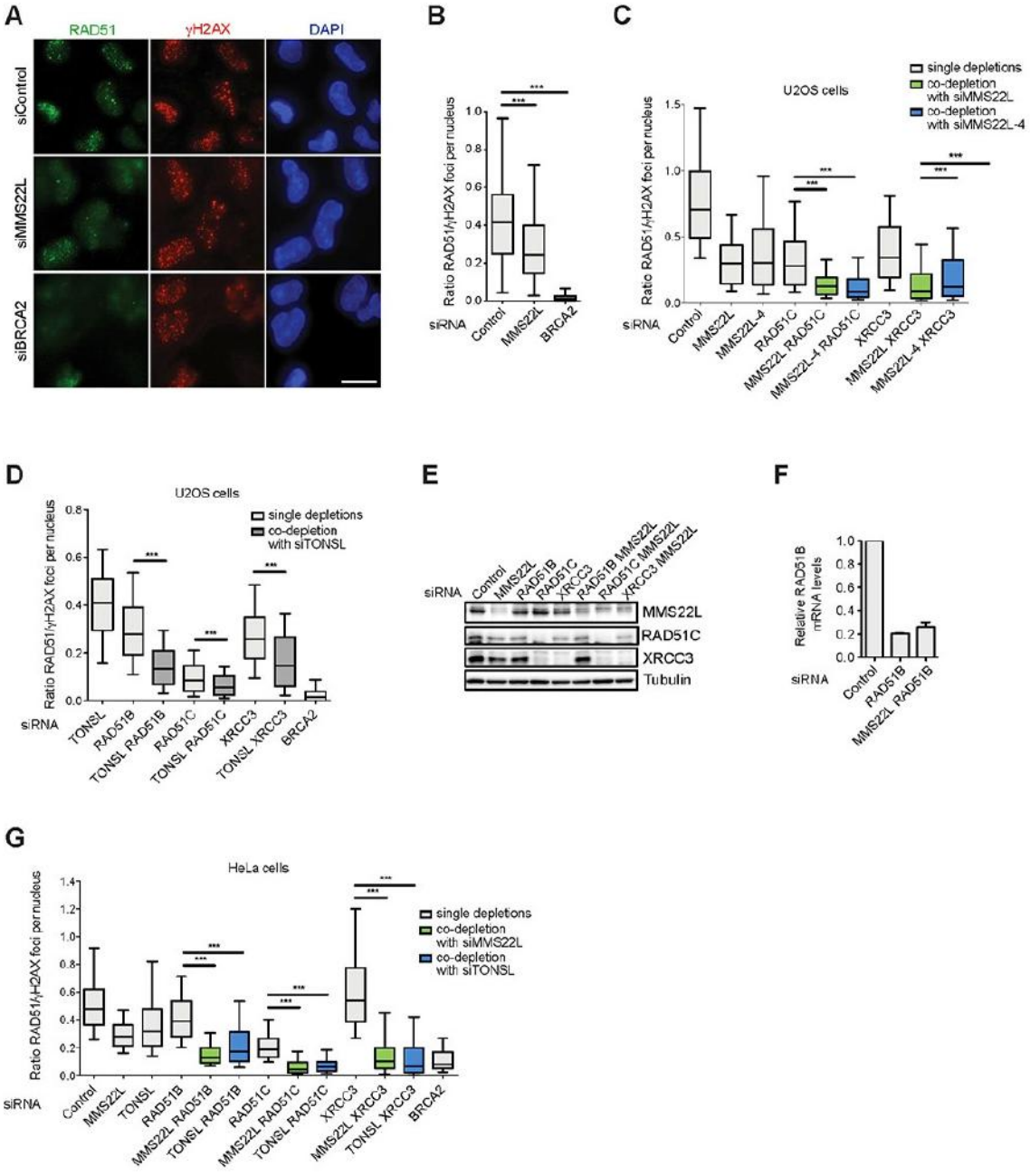
Piwko et al., Figure S4



Piwko et al., Figure S5



Piwko et al., Figure S6



2.2.3 Additional results

2.2.3.1 MBP-tagged RAD51 purification

Successful purification of MBP-tagged RAD51 from bacteria

Introduction: Previous RAD51 purification with intein-tagged RAD51, as described in the manuscript (section 2.2.2), resulted in a relatively low concentration of a final protein prep (~6.7 - 13.5 μM). However, subsequent experiment stated in section 2.2.3.3, required higher concentration of RAD51. Therefore we set out to design new construct with MBP-tagged RAD51 as MBP tag binds amylose resin with higher affinity than intein to chitin beads.

Results and Discussion: We successfully purified two different preps of RAD51, which concentrations were 24 μM and 72 μM (Figure 1 in this chapter, page 93).

2.2.3.2 In vitro binding assay with mitochondrial SSB as a control

MMS22L-TONSL specifically binds RPA-coated ssDNA

Introduction: In the previous section we showed that MMS22L-TONSL binds RPA-coated ssDNA. However, it was not entirely clear, whether such interaction is specific. To test the specificity of the MMS22L-TONSL-RPA interaction, the mitochondrial single-strand binding protein (mtSSB) was pre-bound to biotinylated oligonucleotide and MMS22L-TONSL was added and compared RPA-coated ssDNA.

Results and Discussion: As expected, MMS22L-TONSL bound to RPA-coated ssDNA but did not bind to mtSSB-coated ssDNA. This result clearly shows that MMS22L-TONSL-RPA interaction is specific (Figure 2 in this chapter, page 94).

2.2.3.3 Ternary complex

MMS22L-TONSL and RAD51 bind to RPA-coated ssDNA

Introduction: MMS22L-TONSL binds to RPA-coated ssDNA. Since MMS22L-TONSL is potentially a recombination mediator, the key question was to ask whether MMS22L-TONSL unloads RPA and helps to load RAD51 onto ssDNA filament. The other scenario would be, that MMS22L-TONSL brings RAD51 to RPA-coated ssDNA by forming a ternary complex.

Results and Discussion: We showed that MMS22L-TONSL probably does not load RAD51 onto RPA-coated ssDNA, as the levels of detected RPA were unaffected upon addition of MMS22L-TONSL and RAD51 (Figure 3, lane3, page 95). Nevertheless, we showed that MMS22L-TONSL binds RAD51-coated ssDNA. Furthermore, we also detected MMS22L-TONSL, RPA1/2, and RAD51 in the pull-downs. However, it is not clear, whether RAD51 was bound to RPA-coated ssDNA via MMS22L-TONSL, therefore this is yet to be determined.

2.2.3.4 Mapping of MMS22L-RAD51 and -RPA interacting sites/regions

Successful purification of various MMS22L fragments from Sf9 insect cell system

Introduction: So far, only TONSL subunit domain structure has been described as discussed above. However, no information is provided on structure of MMS22L. In the previous section, we showed that MMS22L subunit alone has the capacity to bind RPA-coated ssDNA (S2J, page 82) and also directly binds RAD51 (ref to figure). We therefore wanted to analyse interaction sites/regions of MMS22L-RPA and MMS22L-RAD51. To do so, we designed 20 MMS22L fragments mapping the entire MMS22L (see Figure 4A in this section and supplementary information in chapter 5.2), which were cloned into pFastBac vector. Since these fragments were very long, we opted to use *Sf9* insect cell system for expression.

Results and Discussion: We successfully purified all fragments. Next, we plan to test them in pull-down assays to identify interaction sites with RPA and RAD51.

Experimental procedures

Electromobility shift assay

The binding reactions were performed as described previously. [101] The binding buffer was composed of 25 mM Tris-HCl (pH 7.5), 2 mM $Mg(OAc)_2$, 1 mM dithiothreitol, and 0.1 mg/ml bovine serum albumin (New England BioLabs). The oligonucleotides were used as listed in the supplementary information 5.1. **ssDNA (50 nt)** was generated by ^{32}P radiolabeling 5' end of X12-3. **dsDNA (50 nt)** was generated by annealing ^{32}P radiolabeled 5' end of X12-3 and X12-4 C in 1:2 molar ratio. **5' tailed substrate (50 nt)** was generated by annealing ^{32}P radiolabeled 5' end of X12-3 and X12-4 SC in 1:2 molar ratio. **3' tailed substrate (50 nt)** was generated by annealing radiolabeled 5' end of X12-3 SC and X12-4 NC in 1:2 molar ratio. **4-way junction (50 nt)** was generated by annealing radiolabeled 5' end of PC 1253 and PC 1254, PC 1255, PC 1256 in 1:2:2:2 molar ratio. **ssDNA (93 nt)** was generated by ^{32}P radiolabeling 5' end of X12-3 TOP L. **dsDNA (93 nt)** was generated by annealing ^{32}P radiolabeled 5' end of X12-3 TOP L and X12-3 BOTTOM LC in 1:2 molar ratio. **2-way substrate (93 nt)** was generated by annealing ^{32}P radiolabeled 5' end of X12-3 TOP L and X12-3 HJ3 in 1:2 molar ratio. **Replication fork (RF) substrate (93 nt)** was generated by annealing ^{32}P radiolabeled 5' end of X12-3 TOP L and X12-3 HJ 1S, X12-3 HJ 2Sb, and X12-3 HJ3 in 1:2:2:2 molar ratio. **4-way junction substrate (93 nt)** was generated by annealing ^{32}P radiolabeled 5' end of X12-3 TOP L and X12-3 HJ1, X12-3 HJ2, and X12-3 HJ3 in 1:2:2:2 molar ratio. **Nicked 4-way junction (n4W) substrate (93 nt)** was generated by annealing ^{32}P radiolabeled 5' end of X12-3 TOP L and X12-3 HJ1, X12-3 HJ2 Sa, X12-3 HJ Sb, X12-3 HJ3 in 1:2:2:2:2 molar ratio. **4WJ-3' tailed substrate (93 nt)** was generated by annealing ^{32}P radiolabeled 5' end of X12-3 TOP L and X12-3 HJ 1, X12-3 HJ2 Sb, and X12-3 HJ3 in 1:2:2:2 molar ratio. **4WJ-5' tailed substrate (93 nt)** was generated by

annealing ^{32}P radiolabeled 5' end of X12-3 TOP L and X12-3 HJ 1S, X12-3 HJ2, and X12-3 HJ3 in 1:2:2:2 molar ratio.

Preparation of recombinant MBP-RAD51 protein

RAD51 was expressed from a pMALT-p vector where it was fused to an MBP-tag at its N-terminus. This vector was generated as follows. RAD51 was amplified from pTXB3 vector by using forward primer: F-R51: 5'CTACTAGGATCCATGGCAATGCAGATGCAGCTT-3' and reverse primer: R-R51: 5'-CTTCTTCTGCAGTCAGTCTTTGGCATCTCCCACTCC-3'. PCR product was digested with BamHI and PstI and cloned into pMALT-p vector. The protein was expressed in *E. coli* BL21(DE3)pLysS cells (Promega) and purified on amylose resin (New England Biolabs) according to manufacturer's recommendation. Briefly, the pellets were lysed with lysis buffer [50 mM Tris-HCl pH 7.5, 1 mM phenylmethanesulfonyl fluoride (PMSF), 1 mM dithithreitol, 10% glycerol, 500 mM NaCl, and protease inhibitors (1;500; Sigma)]. Then, the lysate was sonicated six times at 70% intensity and loaded onto column with pre-equlibrated amylose resin. The immobilized sample was further washed with wash buffer I [50 mM Tris-HCl pH 7.5, 1 mM dithithreitol, 10% glycerol, 1 M NaCl] and wash buffer II [50 mM Tris-HCl pH 7.5, 1 mM dithithreitol, 10% glycerol, 300 mM NaCl]. MBP-RAD51 was eluted with elution buffer [wash buffer II supplemented with 10 mM maltose]. The MBP tag was cleaved with PreScission protease for 2 h at 4 °C. The cleaved RAD51 was diluted with 50 mM Tris-HCl (pH 7.5) to a final concentration of 150 mM NaCl, and loaded onto pre-equilibrated HiTrap Q HP column (GE Healthcare). The column was washed with R buffer containing 20 mM Tris-HCl (pH 7.5), 1 mM EDTA, 0.5 mM dithiothreitol, 10% glycerol supplemented with 150 mM NaCl. RAD51 was eluted with 10 ml gradient of 150 mM to 700 mM NaCl in R buffer supplemented with 1M NaCl. Next day, the final sample was mixed with PreScission protease to cleave the uncleaved MBP tag and dialyzed for 2.5 h against dialysis buffer containing 20 mM Tris-HCl (pH 7.5), 20% glycerol, 100 mM NaCl, 1 mM dithiothreitol. The residual cleaved MBP tag was

captured on amylose resin during 1 h batch-wise incubation. The final sample was then aliquoted and stored at -80 °C.

In vitro binding assay with mtSSB

The assay was done as described previously (refer to the manuscript). Human mtSSB was added to the reaction and treated similarly to the reactions carried out with RPA. Human mtSSB was purified as described in Pinto et al., 2016 (manuscript under review).

Ternary complex assay

The assay was performed similarly as described previously with several changes. [20] Briefly, RPA (100 nM), MMS22L-TONSL (50 nM) and RAD51 (500 nM) were pre-incubated at 37 ºC for 10 min according to a reaction scheme shown in the Figure 3 (chapter 2.2.3) in binding buffer supplemented with ATP (20 mM Tris-HCl (pH 7.5), 120 mM NaCl, 0.1%(v/v) Triton X-100, 2 mM $CaCl_2$, 10 mM $Mg(OAc)_2$, 2 mM ATP, 1 mM dithiothreitol, and 0.1 mg/ml bovine serum albumin) or without ATP (20 mM Tris-HCl (pH 7.5), 120 mM NaCl, 0.1%(v/v) Triton X-100, 2 mM $CaCl_2$, 10 mM $Mg(OAc)_2$, 1 mM dithiothreitol, and 0.1 mg/ml bovine serum albumin) together with biotinylated B-90-mer. Reactions were incubated with pre-equilibrated Dynabeads® M-280 Streptavidin (Invitrogen) rotating at 4 ºC for 30 min. Then the beads were washed four times with wash buffer with or without 2 mM ATP [20 mM Tris-HCl (pH 7.5), 120 mM NaCl, 0.1%(v/v) Triton X-100, 2 mM $CaCl_2$, 10 mM $Mg(OAc)_2$, 1 mM dithiothreitol]. The protein was eluted by boiling with 1x protein loading buffer (50 mM Tris, pH 6.8; 10% glycerol; 1.6% sodiumdodecylsulphate, 0.1 M dithiothreitol; 0.02% bromphenol blue) and 1 µl bovine serum albumin at 95 ºC for 3 min and loaded onto 10 % SDS-polyacrylamide gel. The amounts of RAD51, RPA, and MMS22L-TONSL proteins bound to biotinylated DNA oligonucleotide were estimated by western blotting using a polyclonal RAD51 antibody specific to human RAD51, [102] RPA antibody (Calbiochem), and his tag antibody (GenScript).

Mapping of MMS22L-RAD51 and -RPA interacting regions

Genes coding for MMS22L fragments as listed in supplementary information 5.2 were amplified from pFB-MBP-MMS22L-his by PCR using primers as listed in supplementary information 5.2. The PCR products were digested with ApaI and XhoI restriction endonucleases and cloned into pFastBac1 vector previously digested with ApaI and XhoI, generating pFB-MBP-MMS22L fragment-FLAG.
The MMS22L-FA mutant was prepared from pFB-MMS22L(L1)-FLAG construct. We used Quik Change II XL Site-directed mutagenesis kit (Agilent Technologies) with several modifications. For PCR mix, we used KOD Hot Start DNA polymerase (Novagen) to enhance the reaction efficiency as described previously. [103] The MMS22L fragments were expressed in *Sf9* insect cells by infection with designated MMS22L baculoviruses according to manufacturer's recommendation (Invitrogen). The proteins were extracted and purified similarly as described previously in the manucsript above (chapter 2.2.2) and in Cejka and Kowalczykowski[104] with several changes as decribed below. Briefly, the protein complex or MMS22L alone were first immobilized on amylose resin via MBP tag (2 ml of 50% slurry per MMS22L fragment), then MBP tag was cleaved away by PreScission protease and final MMS22L fragments were aliquoted. Importantly, the NaCl concentration was kept at 250 mM.

Figure legends

Fig. 1. Purified RAD51.

Fig. 2. MMS22L-TONSL does not bind mtSSB-coated ssDNA.

Fig. 3. MMS22L-TONSL binds RAD51 coated ssDNA in the presence of ATP.

Fig. 4. Purified MMS22L fragments from Sf9 insect cell system.

Figure 1

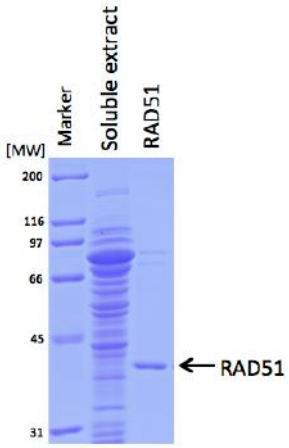


Figure 2

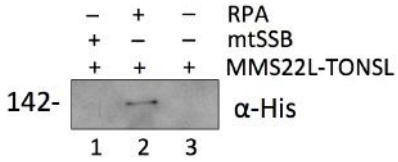


Figure 3

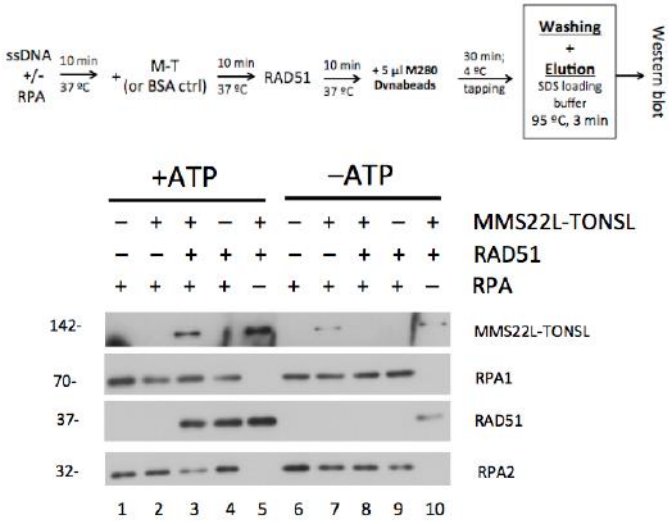
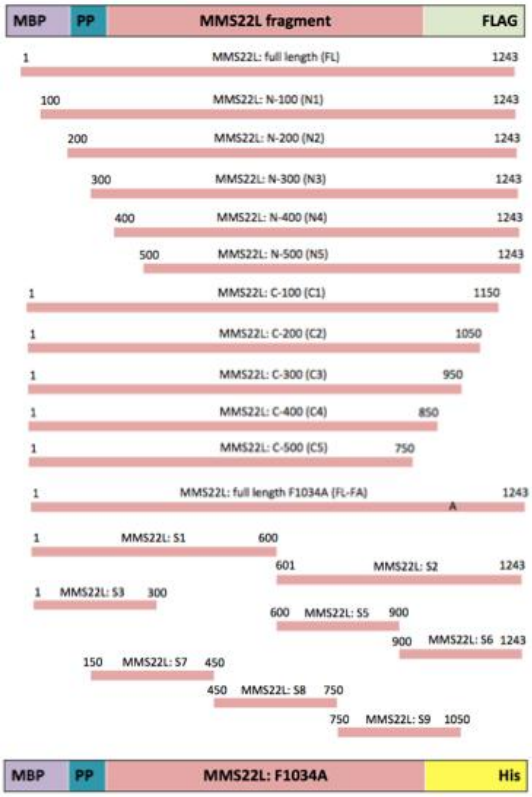
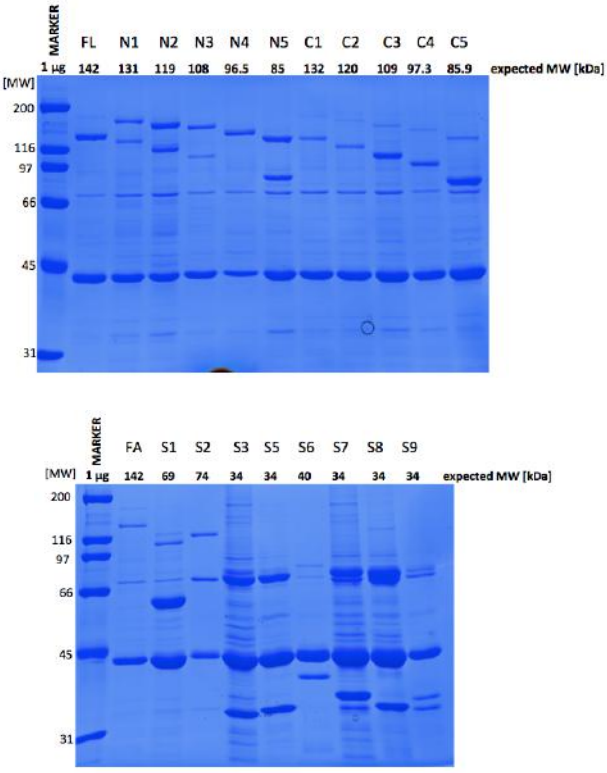

ssDNA +/- RPA
10 min 37 ºC
+ M-T (or BSA ctrl)
10 min 37 ºC
RAD51
10 min 37 ºC
+ 5 µl M280 Dynabeads
30 min; 4 ºC tapping
Washing + Elution SDS loading buffer 95 ºC, 3 min
Western blot
+ATP
−ATP
MMS22L-TONSL
RAD51
RPA
142-
70-
37-
32-
MMS22L-TONSL
RPA1
RAD51
RPA2
1 2 3 4 5 6 7 8 9 10

Figure 4

A


MBP
PP
MMS22L fragment
FLAG
MMS22L: full length (FL)
MMS22L: N-100 (N1)
MMS22L: N-200 (N2)
MMS22L: N-300 (N3)
MMS22L: N-400 (N4)
MMS22L: N-500 (N5)
MMS22L: C-100 (C1)
MMS22L: C-200 (C2)
MMS22L: C-300 (C3)
MMS22L: C-400 (C4)
MMS22L: C-500 (C5)
MMS22L: full length F1034A (FL-FA)
MMS22L: S1
MMS22L: S2
MMS22L: S3
MMS22L: S5
MMS22L: S6
MMS22L: S7
MMS22L: S8
MMS22L: S9
MBP
PP
MMS22L: F1034A
His

B


MARKER
FL N1 N2 N3 N4 N5 C1 C2 C3 C4 C5
1 µg 142 131 119 108 96.5 85 132 120 109 97.3 85.9 expected MW [kDa]
[MW]
200
116
97
66
45
31
MARKER
FA S1 S2 S3 S5 S6 S7 S8 S9
[MW] 1 µg 142 69 74 34 34 40 34 34 34 expected MW [kDa]
200
116
97
66
45
31

2.2.4 H4 K20me0 marks post-replicative chromatin and recruits the TONSL-MMS22L DNA repair complex

Giulia Saredi, Hongda Huang, Colin M. Hammond, Constance Alabert, Simon Bekker-Jensen, Ignasi Forne, Nazaret Reverón-Gómez, Benjamin M. Foster, **Lucie Mlejnkova**, Till Bartke, Petr Cejka, Niels Mailand, Axel Imhof, Dinshaw J. Patel, and Anja Groth.

In print. Nature, June 2016.

I designed the MMS22L-TONSL protein purification protocol together with P.C. and purified the protein complex, which was then used in the pull-down of recombinant TONSL-MMS22L with biotinylated recombinant mononucleosomes shown in the main Figure 2d (page 99).

LETTER

doi:10.1038/nature18312

H4K20me0 marks post-replicative chromatin and recruits the TONSL–MMS22L DNA repair complex

Giulia Saredi[1]*, Hongda Huang[2]*, Colin M. Hammond[1], Constance Alabert[1], Simon Bekker-Jensen[3], Ignasi Forne[4], Nazaret Reverón-Gómez[1], Benjamin M. Foster[5], Lucie Mlejnkova[6], Till Bartke[5], Petr Cejka[6], Niels Mailand[3], Axel Imhof[4], Dinshaw J. Patel[2] & Anja Groth[1]

After DNA replication, chromosomal processes including DNA repair and transcription take place in the context of sister chromatids. While cell cycle regulation can guide these processes globally, mechanisms to distinguish pre- and post-replicative states locally remain unknown. Here we reveal that new histones incorporated during DNA replication provide a signature of post-replicative chromatin, read by the human TONSL–MMS22L[1–4] homologous recombination complex. We identify the TONSL ankyrin repeat domain (ARD) as a reader of histone H4 tails unmethylated at K20 (H4K20me0), which are specific to new histones incorporated during DNA replication and mark post-replicative chromatin until the G2/M phase of the cell cycle. Accordingly, TONSL–MMS22L binds new histones H3–H4 both before and after incorporation into nucleosomes, remaining on replicated chromatin until late G2/M. H4K20me0 recognition is required for TONSL–MMS22L binding to chromatin and accumulation at challenged replication forks and DNA lesions. Consequently, TONSL ARD mutants are toxic, compromising genome stability, cell viability and resistance to replication stress. Together, these data reveal a histone-reader-based mechanism for recognizing the post-replicative state, offering a new angle to understand DNA repair, with the potential for targeted cancer therapy.

The TONSL–MMS22L complex is an obligate heterodimer required for replication fork stability and repair of replication-associated DNA damage by aiding RAD51 loading[1–4]. TONSL–MMS22L associates with soluble non-nucleosomal histones H3–H4 (refs 1, 5), the histone chaperone ASF1 (refs 1–4) and MCM2/4/6/7 (refs 1–5) in a manner that depends on the TONSL ARD[1]. We have found that histones H3–H4 bridge the interactions between TONSL–MMS22L and ASF1 (ref. 1), between ASF1 and MCM2 (refs 6, 7), and between TONSL–MMS22L and MCM2 (Extended Data Fig. 1a), suggesting simultaneous binding of these proteins to histones H3–H4 in a large pre-deposition complex. In addition, TONSL–MMS22L interacts with nucleosomal histones in chromatin (Extended Data Fig. 1b). This suggests that TONSL–MMS22L functions as an H3–H4 histone chaperone while also acting as a histone reader in chromatin. Consistent with a chaperone function, TONSL was recently shown to have histone chaperone activity *in vitro*[5]. We therefore set out to explore the mechanism of action of TONSL–MMS22L by a structure–function approach.

Full-length TONSL and the ARD alone bound directly to histones H3–H4 but not H2A–H2B *in vitro* (Extended Data Fig. 1c–f). As our attempts to crystallize the ARD with H3–H4 were not successful, we linked the ARD to the MCM2 histone-binding domain (HBD), because a similar design previously enabled us to solve the structure of an H3–H4 dimer in complex with MCM2 and ASF1 (ref. 7). We obtained crystals of covalently linked MCM2 HBD–G_4–TONSL ARD in complex with H3 (57–135) and H4 that diffracted to 2.43 Å resolution, and solved the structure by molecular replacement on the basis of our structure of MCM2 HBD in complex with an H3–H4 tetramer[7] (Fig. 1a, b; for X-ray statistics, see Extended Data Table 1). The structure shows a pair of MCM2 HBDs wrapped around the lateral surface of the H3–H4 tetramer, similar to the MCM2-HBD–H3–H4 complex alone[7,8], while two TONSL ARDs interact with each of the H4 tails (Fig. 1a, b). The G_4-linker along with flanking residues formed a 19-residue-long disordered segment that could reach a distance of up to 70 Å. The distance from the observed C terminus of MCM2 HBD to the observed N terminus of TONSL ARD is only 10 Å, indicating that the covalent linkage within the MCM2-HBD–G_4–TONSL-ARD cassette does not affect the structural integrity of the complex. TONSL ARD forms no intermolecular interactions with the MCM2 HBD, consistent with H3–H4 bridging the interaction of TONSL and MCM2 in cells (Extended Data Fig. 1a), and it shows only minimal contacts with the core of the H3–H4 tetramer (Fig. 1a, b). However, the TONSL ARD forms extensive contacts with a segment of the H4 tail (Fig. 1b, c and Extended Data Fig. 1g) and, consistently, it binds the glutathione *S*-transferase (GST)–H4 tail, but not the GST–H3 tail, *in vitro* (Extended Data Fig. 2a). In addition to defining TONSL binding to soluble histone H3–H4 in complex with MCM2 (Fig. 1a, b), this binding mode is also compatible with TONSL binding histone H3–H4 dimers in a co-chaperone complex with MCM2 and ASF1 (ref. 7) as well as recognizing H4 tails in a nucleosome (see models in Extended Data Fig. 2b, c).

The TONSL ARD consists of four ankyrin repeats, three of which adopt the canonical fold (ANK1–3), while the remaining one is an atypical and capping repeat (ANK4) (Extended Data Fig. 1g). The TONSL ARD uses its elongated concave surface to form extensive contacts with the H4 tail in an extended β-strand-like conformation (Extended Data Fig. 1g and Fig. 1c). Notably, 15 out of 18 residues that constitute the H4-tail-binding surface of TONSL ARD are highly conserved (Extended Data Fig. 2d). The TONSL ARD targets the H4 tail, spanning residues Lys12 to Arg23 (Fig. 1c–g and Extended Data Fig. 3a, b) with three consecutive binding channels accommodating Arg17, His18 and Lys20 (Fig. 1d). These H4 residues are part of a basic region, which can interact with the acidic patch on neighbour nucleosomes[9] in compact chromatin. H4 Arg17 forms two hydrogen bonds with ARD Asn571 and stacks with Tyr572 and Cys608 (Fig. 1c, e), while H4 His18 penetrates into a pocket lined by four strictly conserved residues (Trp563, Glu568, Asn571 and Asp604) (Fig. 1c, f). Substitution of H4 His18 with the larger Trp residue (mutant H18W) disrupts

[1]Biotech Research and Innovation Centre (BRIC) and Centre for Epigenetics, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen DK-2200, Denmark. [2]Structural Biology Program, Memorial Sloan-Kettering Cancer Center, New York, New York 10065, USA. [3]The Novo Nordisk Foundation Center for Protein Research, University of Copenhagen, Copenhagen DK-2200, Denmark. [4]Department of Molecular Biology, Biomedical Center and Center for Integrated Protein Science Munich, Ludwig-Maximilians University, 80336 Munich, Germany. [5]MRC Clinical Sciences Centre (CSC) and Institute of Clinical Sciences (ICS), Faculty of Medicine, Imperial College London, Du Cane Road, London W12 0NN, UK. [6]Institute of Molecular Cancer Research, University of Zurich, Zurich CH-8057, Switzerland.
*These authors contributed equally to this work.

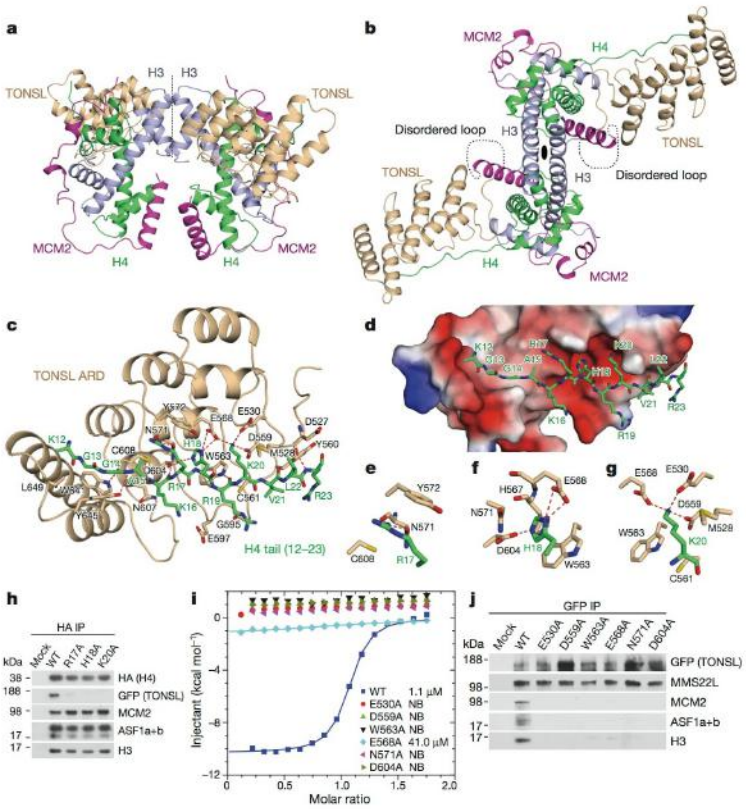


Figure 1 | TONSL ARD interacts with the histone H4 tail. a, **b**, Two different representative views of the overall structure of the TONSL-ARD–MCM2-HBD–H3–H4 tetramer complex. **c**, Intermolecular interactions between TONSL ARD and the H4 tail. **d**, The electrostatic potential surface of ARD showing the acidic concave surface binding site for the H4 tail. **e**–**g**, Highlights of the intermolecular interactions of H4 Arg17, His18 and Lys20 with ARD. **h**, Immunoprecipitation (IP) of soluble haemagglutinin (HA)–SNAP–H4 wild type (WT) or mutant transfected into green fluorescent protein (GFP)–TONSL U-2-OS cells. **i**, ITC of TONSL ARD wild type and mutants with H4 tail peptide. **j**, Immunoprecipitation of soluble GFP–TONSL wild type or mutant. **h**, **j**, Data are representative of three independent experiments. For protein inputs, see Extended Data Fig. 9b, c; for gel source data, see Supplementary Fig. 1.

binding with ARD (Fig. 2a), underscoring the importance of fitting His18 in the pocket. The H4 Lys20 residue is bound within an acidic surface channel on ARD (Fig. 1c, d). The side chain of H4 Lys20 interacts with Met528 and contacts the edge of Trp563 of ARD, while the main-chain atoms of H4 Lys20 packs against Cys561 of ARD (Fig. 1g). The Nζ atom of H4 Lys20 forms three strong hydrogen bonds (distance <3 Å) with the side chains of strictly conserved residues Glu530, Asp559 and Glu568 of ARD, which engage H4 Lys20 in a triangular arrangement (Fig. 1g). Consistent with the structural data, histone H4 mutations R17A, H18A and K20A disrupted binding to TONSL in cells (Fig. 1h). Likewise, mutation of six conserved TONSL residues lining the H4 Arg17, His18 and Lys20 binding channels disrupted binding to H4 peptides and recombinant histone H3–H4 (Fig. 1i and Extended Data Fig. 3c). *In vivo*, these mutants abrogated binding to soluble histone H3–H4 and, consequentially, also association with ASF1a and ASF1b and MCM2 without affecting MMS22L binding to TONSL[1,2] (Fig. 1j and Extended Data Fig. 3d). These mutations did not affect ARD structure, as indicated by circular dichroism (Extended Data Fig. 3e).

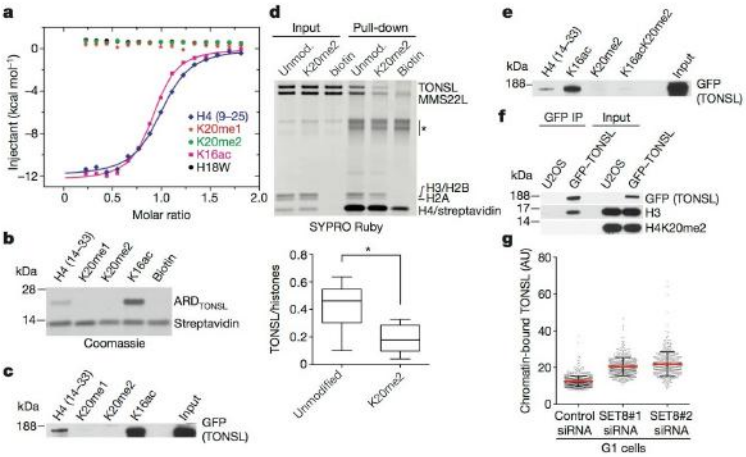


Figure 2 | TONSL ARD recognizes unmodified H4K20. a, ITC of TONSL ARD binding to H4 tail peptides. **b**, **c**, Pull-down of recombinant TONSL ARD (**b**) or GFP–TONSL from extracts (**c**) with biotinylated H4 tail peptides. **d**, Top, pull-down of recombinant TONSL–MMS22L with biotinylated recombinant mononucleosomes. Asterisk indicates unspecific band. Unmod., unmodified. Bottom, TONSL binding quantified relative to histones. Unpaired *t*-test: $*P<0.05$; mean $n=6$; whiskers, outliers. **e**, As in **c**. **f**, Immunoprecipitation of GFP–TONSL from solubilized chromatin. **g**, TONSL chromatin binding in pre-extracted SET8-depleted G1 cells. AU, arbitrary units. Error bars indicate standard deviation (s.d.); $n=747$ (short interfering RNA (siRNA) control), 579 (SET8#1) and 485 (SET8#2). Data are representative of four (**b**), three (**c**) and two (**e**–**g**) independent experiments. For peptide inputs, see Extended Data Fig. 9d.

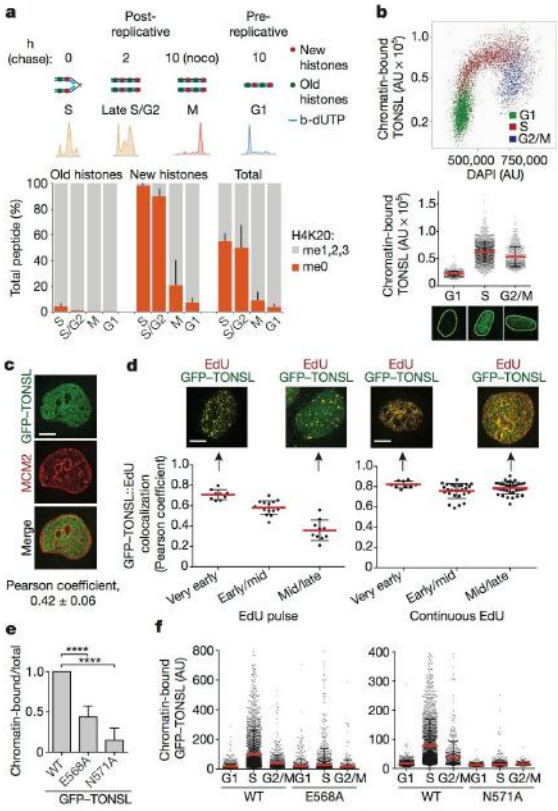


Figure 3 | H4K20me0 is a signature of post-replicative chromatin, read by TONSL ARD. a, H4K20me on new and old histones analysed by stable isotope labelling with amino acids in cell culture (SILAC)-based mass spectrometry of chromatin pulse-labelled with biotin-dUTP (b-dUTP and isolated by NCC (data from ref. 18). noco, nocodazole. Error bars indicate s.d.; $n = 9$ (S), 3 (S/G2, M), 5 (G1); M (old histones) shows the mean of $n = 2$; see also Extended Data Fig. 6a, b. **b**, TONSL chromatin binding in pre-extracted TIG3 fibroblasts shown as a function of 4′,6-diamidino-2-phenylindole (DAPI) intensity or cell cycle stage. Error bars indicate s.d.; $n = 886$ (G1), 2194 (S) and 756 (G2); representative cells are shown. **c**, **d**, Colocalization analysis of chromatin-bound GFP–TONSL with MCM2 (**c**) and EdU (**d**). **d**, Cells were pulsed with EdU (left) or released into S phase in the presence of EdU (right) and analysed by deconvolution microscopy. Error bars indicate s.d.; **c**, $n = 13$; **d**, from left, $n = 9$, 16, 10, 9, 27, 36; representative cells are shown. Scale bar, 5 μm. **e**, Chromatin binding of GFP–TONSL analysed by cellular fractionation, quantified relative to total GFP–TONSL and normalized to wild type (WT). Error bars indicate s.d.; $n = 5$ (wild type/N571A); $n = 3$ (E568A). Unpaired t-test: $^{****}P < 0.0001$. **f**, Chromatin binding of GFP–TONSL analysed as in **b**. Error bars indicate s.d.; from left, $n = 1{,}302$, 3,567, 750, 1,311, 3,850, 838, 1,495, 3,221, 832, 1,695, 2,729, 877. Data are representative of two (**b–d**, **f**) independent experiments.

Together, this defines TONSL ARD as a recognition module for histone H4 tails, distinct from the GLP/G9A ARDs that bind histone H3 tails mono- or dimethylated at K9 (Extended Data Fig. 4a, b)[10].

The structure predicts that methylation on H4K20 should break critical hydrogen bonds with the TONSL ARD. Isothermal titration calorimetry (ITC) and H4-tail peptide pull-downs confirmed that H4K20me1/2 is incompatible with TONSL binding (Fig. 2a–c). Furthermore, H4K20me2 significantly reduced binding of full-length recombinant TONSL–MMS22L to reconstituted mononucleosomes (Fig. 2d). Recently, TONSL ARD with its neighbouring acidic stretch was proposed to bind H3K9me1 (ref. 5), but we were unable to detect an interaction between TONSL ARD (with or without the acidic stretch) and H3K9me1 peptides (Extended Data Fig. 4c, d). Together, our data show that TONSL binds to both free histones and nucleosomes via ARD recognition of H4 tails unmodified at K20 (Figs 1a–j, 2a–e and Extended Data Figs 1b, 2a–c, 3a–d). In line with this, H4K20me2 was not detected on TONSL-bound nucleosomal histones (Fig. 2f), while H4K16ac was present (Extended Data Fig. 5a). H4K16ac stimulated TONSL binding in peptide pull-downs (Fig. 2b, c and Extended Data Fig. 5b) and slightly enhanced ARD binding in ITC (Fig. 2a), but it did not overturn the inhibitory effect of H4K20me2 (Fig. 2e). However, H4K16ac is not essential for TONSL binding *in vivo*, as soluble histone H4 does not carry H4K16ac[11] and depletion of MOF, the major H4K16 acetyltransferase[12], did not significantly affect TONSL binding to chromatin (Extended Data Fig. 5c, d). In contrast, depletion of the H4K20 methyltransferase SET8 (also known as PR-SET7) significantly enhanced TONSL binding to chromatin in G1 cells in which H4K20me2 peaks[13–16] (Fig. 2g and Extended Data Fig. 5e, f). We conclude that TONSL ARD is a histone-reader domain specific for H4 tails unmethylated at K20.

Given that TONSL–MMS22L binds new histones (devoid of H4K20me[11,17]) in a pre-deposition complex with ASF1 and MCM2 (Fig. 1j and Extended Data Fig. 1a)[1], TONSL–MMS22L could be loaded onto replicating DNA together with new histones. To test how long after deposition new histones remain unmethylated at H4K20 with the potential to bind TONSL, we extracted H4K20 data from our recent large-scale proteomic study[18], tracking modifications on new and old recycled histones by nascent chromatin capture (NCC)[19] (Fig. 3a and Extended Data Fig. 6a, b). In nascent chromatin, new histones were exclusively unmethylated at H4K20 (98% H4K20me0), while old recycled histones were almost fully methylated at H4K20 (me1, 7%; me2, 88%; me3, 2%). Consistent with previous work[13–16], our analysis of primary cells (Extended Data Fig. 6c) and degradation of SET8 in S phase[15,16], new histones became methylated in late G2/M, rendering G1 chromatin devoid of H4K20me0 (Fig. 3a). This identifies H4K20me0 on new histones as a signature of post-replicative chromatin, implying that TONSL–MMS22L can bind H4 tails on new histones at replication forks and sister chromatids until late G2/M. Confirming this prediction, TONSL accumulated on chromatin in S phase, remained chromatin-bound in a population of G2 cells, and was excluded from chromatin in G1 (Fig. 3b and Extended Data Fig. 6d–f). To discriminate pre- and post-replicative chromatin, we labelled replicating DNA with 5-ethynyldeoxyuridine (EdU; pulse to mark ongoing replication, continuous labelling to identify post-replicative chromatin) and stained pre-replicative chromatin with MCM2 (refs 20, 21), and analysed colocalization with TONSL. TONSL staining was mutually exclusive with MCM2 (Fig. 3c and Extended Data Fig. 7a), but colocalized with EdU pulse labelling in very early S phase and with replicated DNA (continuous EdU labelling) throughout S phase (Fig. 3d and Extended Data Fig. 7b, c). TONSL was present at sites of ongoing DNA replication throughout S phase, but the degree of colocalization declined in mid/late S phase (Fig. 3d, left), consistent with TONSL binding to post-replicative chromatin also after fork passage (Fig. 3d, right). Mutation of the TONSL ARD abrogated recruitment of TONSL to chromatin, including DNA replication sites (Fig. 3e, f and Extended Data Fig. 7d–g). Together, these data demonstrate that TONSL is recruited to replication forks and post-replicative chromatin via ARD recognition of H4K20me0 on new histones.

Mutation of TONSL ARD also abrogated chromatin binding and recruitment to replication forks in the presence of replication poisons such as camptothecin (CPT) and hydroxyurea (Fig. 4a–c). Furthermore, ARD mutation prevented accumulation of TONSL at site-specific double-strand breaks (DSBs; Fig. 4d and Extended Data Fig. 8a) and microlaser-generated DNA damage (Fig. 4e and Extended Data Fig. 8b, c). Co-staining with cell cycle markers confirmed that

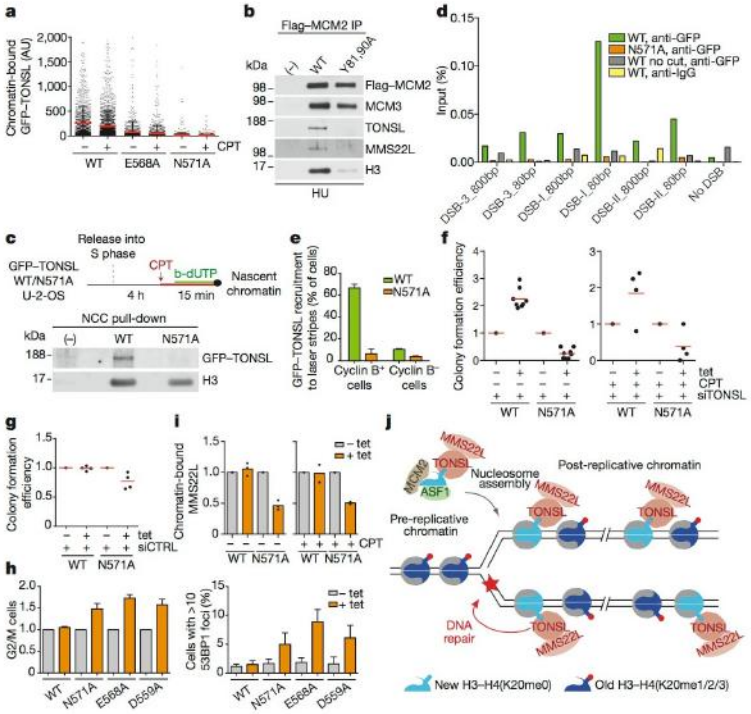


Figure 4 | H4K20me0 recognition is required for TONSL accumulation at DNA repair sites and genome stability. a, Chromatin-binding of GFP–TONSL in CPT-treated S-phase cells. WT, wild type. Error bars indicate s.d.; from left, $n = 1,461$, 2,631, 1,245, 1,764, 2,116, 3,178. **b**, Co-immunoprecipitation (IP) of TONSL–MMS22L with Flag–HA–MCM2 wild type or histone-binding mutant (Y81A, Y90A)[7] from chromatin after hydroxyurea (HU) treatment. **c**, NCC analysis of GFP–TONSL recruitment to replication forks in CPT-treated cells. Minus sign indicates no b-dUTP. **d**, Chromatin immunoprecipitation (ChIP) and quantitative polymerase chain reaction (qPCR) analysis of GFP–TONSL recruitment to site-specific DSBs induced by AsiSI[29]. See Extended Data Fig. 8a for additional controls. **e**, GFP–TONSL recruitment to laser-induced DNA lesions (error bars indicate s.d.; $n = 3$; total cells counted, 210 (wild type) and 252 (N571)). **f**, **g**, Colony formation upon GFP–TONSL induction by tetracycline (+tet) in siRNA- and CPT-treated cells. **h**, Cell cycle and 53BP1 foci analysed by microscopy. Left, percentage G2/M cells shown relative to non-induced (−tet). Error bars indicate s.d., $n = 4$ (left), 5 (right). **i**, Chromatin-bound MMS22L analysed as in Fig. 3e. Mean with individual data points are shown ($n = 3$ (untreated), 2 (CPT)), see Extended Data Fig. 8i for western blots. **j**, TONSL–MMS22L identifies post-replicative chromatin by binding H4K20me0 on new histones, directing TONSL–MMS22L genome surveillance function to DNA having a sister chromatid. Data are representative of three (**a**), two (**b**–**d**, **f**, right, **g**), and four (**f**, left) independent experiments. For protein inputs, see Extended Data Fig. 9e, f.

TONSL is recruited to DNA repair sites only in S and G2 cells, as expected[2] (Fig. 4e and Extended Data Fig. 8d, e). We conclude that H4K20me0 binding is required for TONSL accumulation at damaged forks and DNA lesions in post-replicative chromatin. However, this was not due to increased H4K20me0 (Extended Data Fig. 8f), suggesting that unmasking of H4 tails upon chromatin decompaction[9,22] and/or interaction with repair factors contribute to TONSL–MMS22L accumulation at repair sites. Consistent with an auxiliary mode of recruitment, MMS22L interaction with RAD51 can stabilize the complex at challenged forks (P. Cejka and M. Peter, personal communication). Our data suggest that this is subsequent to H4K20me0 binding (Fig. 4a–e), and we thus next addressed the contribution of H4K20me0 recognition to TONSL–MMS22L function. In complementation analysis, TONSL wild type partially rescued the viability of TONSL-depleted cells in the presence and absence of CPT (Fig. 4f and Extended Data Fig. 8g, h), whereas TONSL ARD mutants were highly toxic (Fig. 4f and Extended Data Fig. 8g, h). In control cells, TONSL ARD mutants also reduced viability, causing G2/M arrest accompanied by replication-associated DNA damage (Fig. 4g, h). Furthermore, TONSL ARD mutants titrated MMS22L away from chromatin (Fig. 4i and Extended Data Fig. 8i), explaining the dominant-negative phenotype that mimics TONSL–MMS22L depletion[1–4]. Collectively, this indicates that recognition of H4K20me0 is central to TONSL–MMS22L function in safeguarding genome stability.

This study reveals that post-replicative chromatin has a distinct histone modification signature, read by the TONSL–MMS22L effector protein (Fig. 4j). This opens a new avenue to understand how DNA repair and other chromosomal transactions can be directly linked to the replication state of a genomic locus. Intriguingly, it is the new histones that make post-replicative chromatin distinct, and in this way H4K20me0 resembles the behaviour of H3K56ac[23] in yeast. Our data indicate that TONSL–MMS22L is delivered to nascent chromatin with new histones via the pre-deposition complex with MCM2 and ASF1 (Fig. 4j). We favour the idea that TONSL has a dual function as a histone chaperone[5] and histone reader. Our structural work proposes that TONSL acts in a histone chaperone-like capacity by sequestering the H4 tail to prevent spurious contacts with DNA during H3–H4 deposition. Furthermore, TONSL ARD may counteract chromatin compaction by preventing association of the H4 tail with the H2A–H2B acidic patch on neighbouring nucleosomes. Thus, TONSL changes our perception of a histone chaperone by binding both soluble and nucleosomal histones. In its function as a histone reader, TONSL localizes MMS22L to post-replicative chromatin via H4K20me0 and allows TONSL–MMS22L to accumulate at damaged forks and DNA

lesions. We envision that H4K20me0 works as an affinity trap, making TONSL–MMS22L readily available to support RAD51 loading during homologous recombination. This provides a new approach and opportunity to understand the role of H4K20 in DNA repair, complementing the well-described role of H4K20me1/2 in recruiting 53BP1 to promote non-homologous end joining in competition with BRCA1–BARD1 (refs 24, 25). In post-replicative chromatin, H4K20me1/2 on old histones will support 53BP1 recruitment. Whether H4K20me0 on new histones also influences DNA repair pathway choice will be of interest in future investigations. It is notable that the structure of the TONSL ARD, including the histone-binding surface, is highly similar to the ARD of BARD1 (Extended Data Fig. 9a)[26], required for BRCA1 tumour suppressor function and homologous recombination[27]. Multiple mutations in the TONSL ARD are reported in cancer (C608G, P557S, E597K; http://www.cancer.sanger.ac.uk) and the N571 residue, key to histone H4 binding, corresponds to the BARD1 N470S cancer mutation[26,28]. This highlights the tumour suppressor function of H4K20me0 recognition, and the possibilities it brings for targeted cancer therapy should be explored in the future.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 14 August 2015; accepted 5 May 2016.
Published online xx xx 2016.

1. Duro, E. *et al.* Identification of the MMS22L-TONSL complex that promotes homologous recombination. *Mol. Cell* **40,** 632–644 (2010).
2. O'Donnell, L. *et al.* The MMS22L-TONSL complex mediates recovery from replication stress and homologous recombination. *Mol. Cell* **40,** 619–631 (2010).
3. O'Connell, B. C. *et al.* A genome-wide camptothecin sensitivity screen identifies a mammalian MMS22L-NFKBIL2 complex required for genomic stability. *Mol. Cell* **40,** 645–657 (2010).
4. Piwko, W. *et al.* RNAi-based screening identifies the Mms22L–Nfkbil2 complex as a novel regulator of DNA replication in human cells. *EMBO J.* **29,** 4210–4222 (2010).
5. Campos, E. I. *et al.* Analysis of the histone H3.1 interactome: a suitable chaperone for the right event. *Mol. Cell* **60,** 697–709 (2015).
6. Groth, A. *et al.* Regulation of replication fork progression through histone supply and demand. *Science* **318,** 1928–1931 (2007).
7. Huang, H. *et al.* A unique binding mode enables MCM2 to chaperone histones H3–H4 at replication forks. *Nature Struct. Mol. Biol.* **22,** 618–626 (2015).
8. Richet, N. *et al.* Structural insight into how the human helicase subunit MCM2 may act as a histone chaperone together with ASF1 at the replication fork. *Nucleic Acids Res.* **43,** 1905–1917 (2015).
9. Kalashnikova, A. A., Porter-Goff, M. E., Muthurajan, U. M., Luger, K. & Hansen, J. C. The role of the nucleosome acidic patch in modulating higher order chromatin structure. *J. R. Soc. Interface* **10,** 20121022 (2013).
10. Collins, R. E. *et al.* The ankyrin repeats of G9a and GLP histone methyltransferases are mono- and dimethyllysine binding modules. *Nature Struct. Mol. Biol.* **15,** 245–250 (2008).
11. Jasencakova, Z. *et al.* Replication stress interferes with histone recycling and predeposition marking of new histones. *Mol. Cell* **37,** 736–743 (2010).
12. Taipale, M. *et al.* hMOF histone acetyltransferase is required for histone H4 lysine 16 acetylation in mammalian cells. *Mol. Cell. Biol.* **25,** 6798–6810 (2005).
13. Rice, J. C. *et al.* Mitotic-specific methylation of histone H4 Lys 20 follows increased PR-Set7 expression and its localization to mitotic chromosomes. *Genes Dev.* **16,** 2225–2230 (2002).
14. Pesavento, J. J., Yang, H. & Kelleher, N. L. & Mizzen, C. A. Certain and progressive methylation of histone H4 at lysine 20 during the cell cycle. *Mol. Cell. Biol.* **28,** 468–486 (2008).
15. Beck, D. B., Oda, H., Shen, S. S. & Reinberg, D. PR-Set7 and H4K20me1: at the crossroads of genome integrity, cell cycle, chromosome condensation, and transcription. *Genes Dev.* **26,** 325–337 (2012).
16. Jørgensen, S., Schotta, G. & Sørensen, C. S. Histone H4 lysine 20 methylation: key player in epigenetic regulation of genomic integrity. *Nucleic Acids Res.* **41,** 2797–2806 (2013).
17. Loyola, A., Bonaldi, T., Roche, D., Imhof, A. & Almouzni, G. PTMs on H3 variants before chromatin assembly potentiate their final epigenetic state. *Mol. Cell* **24,** 309–316 (2006).
18. Alabert, C. *et al.* Two distinct modes for propagation of histone PTMs across the cell cycle. *Genes Dev.* **29,** 585–590 (2015).
19. Alabert, C. *et al.* Nascent chromatin capture proteomics determines chromatin dynamics during DNA replication and identifies unknown fork components. *Nature Cell Biol.* **16,** 281–293 (2014).
20. Prasanth, S. G., Méndez, J., Prasanth, K. V. & Stillman, B. Dynamics of pre-replication complex proteins during the cell division cycle. *Phil. Trans. R. Soc. B* **359,** 7–16 (2004).
21. Takahashi, T. S., Wigley, D. B. & Walter, J. C. Pumps, paradoxes and ploughshares: mechanism of the MCM2–7 DNA helicase. *Trends Biochem. Sci.* **30,** 437–444 (2005).
22. Seeber, A., Hauer, M. & Gasser, S. M. Nucleosome remodelers in double-strand break repair. *Curr. Opin. Genet. Dev.* **23,** 174–184 (2013).
23. Burgess, R. J. & Zhang, Z. Histone chaperones in nucleosome assembly and human disease. *Nature Struct. Mol. Biol.* **20,** 14–22 (2013).
24. Botuyan, M. V. *et al.* Structural basis for the methylation state-specific recognition of histone H4–K20 by 53BP1 and Crb2 in DNA repair. *Cell* **127,** 1361–1373 (2006).
25. Panier, S. & Boulton, S. J. Double-strand break repair: 53BP1 comes into focus. *Nature Rev. Mol. Cell Biol.* **15,** 7–18 (2014).
26. Fox, D., III *et al.* Crystal structure of the BARD1 ankyrin repeat domain and its functional consequences. *J. Biol. Chem.* **283,** 21179–21186 (2008).
27. Laufer, M. *et al.* Structural requirements for the BARD1 tumor suppressor in chromosomal stability and homology-directed DNA repair. *J. Biol. Chem.* **282,** 34325–34333 (2007).
28. Ishitobi, M. *et al.* Mutational analysis of BARD1 in familial breast cancer patients in Japan. *Cancer Lett.* **200,** 1–7 (2003).
29. Iacovoni, J. S. *et al.* High-resolution profiling of γH2AX around DNA double strand breaks in the mammalian genome. *EMBO J.* **29,** 1446–1457 (2010).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank the beam staff at the synchrotrons at the Argonne National Laboratory (NE-CAT) for technical assistance. We thank J. Rouse, D. Durocher, G. Legube and C. Storgaard Sørensen for reagents, G. Montoya for assistance with circular dichroism, C. B. Strømme, A. Strandsby, K. Nakamura, S.-b. Lee and M. Hödl for help with experiments, and Y. Antoku for assistance with microscopy. We thank J. Lukas for comments on the manuscript and Z. Jasencakova for illustrations. G.S. was supported by European Commission Marie Curie ITN FP7 'aDDRess'. D.J.P. was supported in part by grants from the Leukemia and Lymphoma Society and the STARR foundation. A.G. is an EMBO Young Investigator and her research is supported by the European Research Council (ERC StG, no. 281765), the Danish National Research Foundation to the Center for Epigenetics (DNRF82), the Danish Cancer Society, the Danish Medical Research Council, the Novo Nordisk Foundation and the Lundbeck Foundation. A.I. is supported by the European Commission FP7 Network of Excellence EpiGeneSys (project 257082), the DFG Excellence Clusters CIPSM and SyNergy, as well as the DFG Collaborative Research Center 1064 (projects A3 and Z3). T.B. is supported by the Medical Research Council and the European Research Council (ERC StG, no. 309952).

Author Contributions G.S. and A.G. conceived and led the functional studies. H.H conceived and led the generation of cassette to crystallize the complex, and H.H. solved the structure and performed the ITC under the supervision of D.J.P. C.M.H. performed peptide pull-downs with ARD and recombinant nucleosomes. C.A. performed SET8 experiments and NCC. S.B.-J. and N.M. analysed recruitment to laser-induced DNA damage. N.R.-G. prepared histones for mass spectrometry and performed ChIP analysis, I.F. analysed histone modifications by mass spectrometry under the supervision of A.I. B.M.F and T.B. prepared modified recombinant nucleosomes. L.M. and P.C. prepared recombinant TONSL–MMS22L. G.S., H.H., D.J.P. and A.G. wrote the manuscript and all authors commented on the manuscript.

Author Information Coordinate and structure factors have been deposited in the Protein Data Bank under accession number 5JA4. Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to D.J.P. (pateld@mskcc.org) or A.G. (anja.groth@bric.ku.dk).

Reviewer Information *Nature* thanks T. Kutateladze and the other anonymous reviewer(s) for their contribution to the peer review of this work.

METHODS

Protein expression and purification. No statistical methods were used to predetermine sample size. The experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment. All proteins used in this study, unless otherwise indicated, were expressed in BL21(DE3)-RIL cell strain (Stratagene). The human TONSL ARD (residues 512–692) and MCM2 HBD (fragments 61–130) were covalently linked through a four-glycine linker (G_4 linker) into one expression cassette. The MCM2 HBD–G_4–TONSL ARD expression cassette was cloned into a modified RSFDuet-1 vector (Novagen), with an N-terminal His_6-SUMO tag. The resulting plasmid was co-expressed with plasmid harbouring histone genes H3.3(Δ56) and H4. The expressed protein complex was first purified on an Ni-NTA affinity column. After removing the His_6-SUMO tag by using Ulp1 (SUMO protease), the protein complex was further purified on HiLoad 16/600 Superdex 200 column (GE Healthcare).

The GST-tagged TONSL ARD and its mutants including E530A, D559A, W563A, E568A, N571A and D604A were cloned into pGEX-6P-1 vector (GE Healthcare). The expressed proteins were first purified using Glutathione Sepharose 4B, then further purified by gel-filtration step. In some case, the GST tag was removed with 3C protease before the gel-filtration step. For purification of GST–H3 tail and GST–H4 tail proteins, the human histones H3 fragment 1–59 and H4 fragment 1–31 were cloned into pGEX-6P-1 vector respectively. The proteins were expressed and purified in the same way.

For production of recombinant full-length TONSL–MMS22L heterodimer, the sequence coding for full-length MMS22L was fused with an MBP tag at the 5′ end and 10× His tag at the 3′ end. The sequence coding for full-length TONSL was fused with a GST tag at the 5′ end. Both MMS22L and TONSL constructs were cloned into a pFastBac1 vector. The complex was expressed in Sf9 cells by co-infection with both recombinant baculoviruses according to manufacturer's recommendation (Invitrogen). The proteins were extracted from Sf9 cells and purified similarly as described previously for Sgs1 (ref. 30). Briefly, the complex was purified on amylose resin, and MBP and GST tags were subsequently cleaved with PreScission protease. The heterodimer was then further purified using a Ni-NTA affinity resin. Washes were performed with 300 mM NaCl buffer.

Crystallization. At first, we tried to crystallize TONSL ARD in complex with a H4 tail or H3–H4 tetramer, but failed even with extensive screening. An additional binding protein may help to stabilize the whole complex and help crystallization. Then we tried to crystallize TONSL ARD in complex with the MCM2 HBD and H3–H4 tetramer. We just got very tiny crystals for this complex, but failed to get big and well-diffracted crystals. We realized that the whole complex of TONSL ARD with MCM2 HBD and H3–H4 tetramer might be destabilized by the harsh crystallization conditions and form a subcomplex, thus hindering the optimization of the crystals. Then we tried to covalently link TONSL ARD and MCM2 HBD into one cassette through different length of glycine linker (G_n linker). The G_{12}, G_{11}, G_{10}, G_9, G_8, G_7, G_6, G_5 and G_4 linkers had been tried and all these cassettes could be crystallized. One of the constructs with a G_4 linker gave well-diffracted crystals.

The G4 linker complex, MCM2-HBD–G_4–TONSL-ARD cassette–H3.3(Δ56)–H4 complex (herein denoted as TONSL-ARD–MCM2-HBD–H3–H4 tetramer complex) at a concentration of 23 mg ml^{-1} was crystallized in 0.1 M MES pH 5.6, 7% isopropanol using sitting-drop vapour-diffusion method at 20 °C. All the crystals were soaked in a cryoprotectant made from mother liquor supplemented with 25% glycerol before flash freezing in liquid nitrogen.

Structure determination. The data sets for the TONSL-ARD–MCM2-HBD–H3–H4 tetramer complex were collected at 0.979 Å on 24-ID-C/E NE-CAT (Advanced Photo Source, Argonne National Laboratory). All the data sets were processed by using the HKL 2000 program. The initial structure for the TONSL-ARD–MCM2-HBD–H3–H4 tetramer complex was solved by molecular replacement in PHASER[31] with our previous structure of the MCM2-HBD–H3–H4 tetramer complex[7] as a search model and manually refined and built using Coot[32]. The final structure of this complex was refined to 2.43 Å resolution using PHENIX[33]. The Ramachandran plot showed 95.9% favoured and 4.1% allowed. Extended Data Table 1 summarizes the statistics for data collection and structural refinement.

Preparation of recombinant modified mononucleosomes. Recombinant human histone proteins were expressed in *Escherichia coli* BL21(DE3)-RIL cells from pET21b(+) (Novagen) vectors and purified by denaturing gel filtration and ion-exchange chromatography essentially as described[34]. All histone proteins were dialysed into water containing 1 mM dithiothreitol (DTT), lyophilized and stored dry at −80 °C. Modified H4 proteins were generated by native chemical ligation essentially as described for H3 (ref. 35). Briefly, tail-less H4 Δ1–28 I29C protein was expressed in *E. coli* BL21(DE3)-RIL cells from pET24b(+) (Novagen) and purified by denaturing gel filtration and reversed-phase chromatography using a ResourceRPC column (GE Healthcare). Purified H4 Δ1–28 I29C was then ligated to N-terminally acetylated H4 1–28 thioester peptides (Almac) and full-length ligated H4 was separated from unligated H4 Δ1–28 I29C by reversed-phase chromatography via a C18 column (Aquapore RP-300/Perkin Elmer) using a gradient from 35% B to 45% B over 20 column volumes (A: 0.1% TFA in water; B: 90% acetonitrile; 0.1% TFA). Ligated H4 was directly lyophilized and stored dry at −80 °C. Ligated H4 was refolded into octamers together with purified histones H2A, H2B and H3.1 and then assembled into nucleosomes with biotinylated 601-DNA as described[34,35].

GST pull-downs. For pull-downs of GST–ARD and its mutants including E530A, D559A, W563A, E568A, N571A and D604A with H3–H4, first 25 μl of Glutathione Sepharose 4B beads were suspended with 200 μl of binding buffer (20 mM Tris pH 7.5 and 0.5 M NaCl), and 1 nmol of GST–ARD proteins were added and incubated at 23 °C for 10 min; then 0.5 nmol of pre-purified H3/H4 tetramers were added and incubated for another 1 h; then the beads were washed quickly with five times 1 ml of washing buffer (binding buffer, 1% Triton X-100) before adding 50 μl of sample loading buffer. An aliquot of 20 μl of each sample was analysed with SDS–PAGE. The GST pull-downs of histone tails of GST–$H3_{1-59}$ and GST–$H4_{1-31}$ with TONSL ARD were performed similarly.

Circular dichroism. Circular dichroism spectra were acquired using a Jasco J-815 Circular Dichroism Spectropolarimeter with a 1 mm quartz cuvette. Spectra were recorded for wild-type and mutant TONSL ARD (512–692, 6.25 μM) between 260 nm and 195 nm in KH2PO4/K2HPO4 buffer (25 mM, pH 7.8) with a data pitch of 0.5 nm, bandwidth of 1 nm and with three accumulations at a scanning speed of 50 nm min^{-1}.

***In vitro* translation and pull-downs with H3–H4 sepharose beads.** NHS-activated sepharose 4 Fast-Flow beads (GE Healthcare) were washed with 0.1 M HCl and incubated overnight with 1 μM recombinant histone H3.1–H4 tetramers (New England Biolabs, catalogue number M2509S) or 1 μM recombinant histone H2A–H2B dimers (New England Biolabs, catalogue number M2508S) in Coupling Buffer (0.2 M $NaHCO_3$, 0.2 M NaCl). One microgram of pSC-B-TONSL, pEXPR-IBA-105-ASF1A wild-type and pEXPR-IBA-105-ASF1A V94R plasmids was incubated with TnT Quick Coupled Transcription/Translation System (Promega) and ^{35}S-methionine according to the manufacturer's instructions. Ten microlitres of *in vitro* translation (IVT) mixture were added to the H3.1–H4, or H2A–H2B, sepharose beads and incubated for 2 h. Beads were washed with 200 mM NaCl, 0.2% NP40 buffer. Beads were boiled in 1× LSB and loaded on a 4–12% Bis-Tris NuPage gel (LifeTechnologies). Proteins were transferred to a 0.2 μm nitrocellulose membrane by overnight wet transfer at 20 V and the membrane was incubated in an autoradiography cassette for 24 h before detection by Phosphor Imager (PerkinElmer).

ITC experiments. All the ITC titrations were performed on a Microcal ITC 200 calorimeter at 25 °C or 20 °C. The peptides of H4 (residues 9–25) and its modified peptides K16ac (with acetylation on Lys16), H18W (with His18 mutated to Trp18), H4K20me1 (monomethylation on Lys20) and H4K20me2 (dimethylation on Lys20), and peptide of H3(1–21)K9me1 (monomethylation on Lys9) were all synthesized at Tufts University Core Facility The exothermic heat of the reaction was measured by 17 sequential 2.2 μl injections of the peptides (1.41 mM in buffer 20 mM Tris pH 7.5 and 0.5 M NaCl) into 200 μl of the TONSL ARD solution (145 μM in the same buffer), spaced at intervals of 150 s or 180 s. The data were processed with Microcal Origin software and the curves were fit to a single site binding model.

Peptide pull-downs assays. Purified recombinant TONSL ARD (residues 512–692) was stored at 400 μM in 1 M NaCl, 20 mM Tris HCl pH 7.5 at −80 °C. For each pull-down, 400 pmol of the ARD stock (1 μl, 400 μM) was diluted with 99 μl of binding buffer (150 mM NaCl, 50 mM Tris HCl pH 7.5, 5% glycerol, 0.25% NP-40, 0.2 mM EDTA, 0.5 mM DTT, 0.2 mM PMSF, 1 mM leupeptin, 1 mM pepstatin). ARD input material was scaled to the number of pull-downs performed. For each pull-down, an H4 peptide (JPT Peptide Technologies GmbH) spanning residues 14–33 (2.5 μl, 250 μM) with a C-terminal biotinoyl-lysine residue or, as control, biotin (2.5 μl, 400 μM) was added to 1.1 ml of binding buffer in addition to 100 μl of the ARD input material and the mixture was incubated overnight rotating at 4 °C. The next day, 25 μl of MyOne Streptavidin C1 beads (Life Technologies) was washed in binding buffer (3 × 500 μl) for each pull-down, removing the final wash from the beads. The ARD/peptide or ARD/biotin mixture was added to an aliquot of pre-washed MyOne Streptavidin C1 beads and incubated with rotation at 4 °C for 3 h. Finally, the beads were washed (2 × 300 μl and 1 × 200 μl of 300 mM NaCl, 50 mM Tris HCl pH 7.5, 5% glycerol, 0.25% NP-40, 0.2 mM EDTA, 0.5 mM DTT, 0.2 mM PMSF, 1 mM leupeptin, 1 mM pepstatin) and pull-down material was visualized by Coomassie staining after SDS–PAGE separation of proteins on a NuPAGE 4–12% gel.

For pull-downs from cell extracts, MyOne T1 beads were incubated overnight with 1 μg of biotinylated peptides in high salt (HS; 300 mM NaCl, 0.5% NP40, Tris HCl, EDTA, 5% glycerol) buffer and subsequently washed twice with PBS. One milligram of NP40/NaCl extract from HeLa S3 or GFP–TONSL U-2-OS cells was added to the beads and incubated for 2 h rotating at 4 °C. The beads were

then washed five times with HS buffer, 2 min rotating at 4 °C. After washing, the beads were resuspended in 1× LSB and boiled for 10 min. The eluted proteins were loaded on a 4–12% Bis-Tris NuPage gel (Life Technologies). Proteins were then transferred to a 0.2 μm nitrocellulose membrane by overnight wet transfer at 20 V and detected by western blotting.

Nucleosome pull-down assay. Modified nucleosomes of H4K20me0 or H4K20me2 were prepared by peptide ligation and stored at 0.1 μg μl^{-1} (by histone octamer) in 100 mM NaCl, 50 mM Tris HCl pH 7.5 at 4 °C. Full-length TONSL–MMS22L complex was stored at 746 nM in 100 mM NaCl, 50 mM Tris HCl, pH 7.5, 5 mM β-mercaptoethanol, 10% glycerol, 0.5 mM PMSF at −80 °C. Nucleosome pull-downs were performed across two sets of conditions ($n = 3$ for each condition) in the presence of herring sperm competitor DNA. Condition number 1: Nucleosomes (1 μg by histone octamer) or biotin (0.5 μg) were mixed with TONSL–MMS22L (1.9 pmol) and made up to 30 μl with binding buffer (500 mM NaCl, 50 mM Tris HCl pH 7.5, 20% glycerol, 0.1% NP-40, 1 mM DTT, protease inhibitors and 10 μg ml^{-1} herring sperm DNA (Sigma)). Inputs of 10 μl were taken before diluting each sample with binding buffer to a final volume of 300 μl and incubating overnight at 4 °C. Pull-downs were performed by adding 20 μl of MyOne Streptavidin C1 beads prewashed and resuspended in 100 μl of binding buffer to each pull-down reaction, incubating at 4 °C for 2 h, washing with 5× 500 μl binding buffer for 2 min at room temperature. Condition number 2: nucleosomes (0.5 μg by histone octamer) or biotin (0.5 μg) were mixed with TONSL–MMS22L (1.3 pmol) and made up to 30 μl with binding buffer (500 mM NaCl, 50 mM Tris HCl pH 7.5, 5% glycerol, 0.5% NP-40, 0.2 mM EDTA, 1 mM DTT, protease inhibitors and 10 μg ml^{-1} herring sperm DNA (Sigma)). Inputs of 15 μl were taken before diluting each sample with binding buffer to a final volume of 500 μl and incubating overnight at 4 °C. Pull-downs were performed by adding 10 μl of MyOne Streptavidin T1 beads prewashed and resuspended in 100 μl of binding buffer to each pull-down reaction, incubating at 4 °C for 4 h, washing with 5 × 500 μl binding buffer for 2 min at room temperature. Pull-downs were visualized by SYPRO Ruby staining after SDS–PAGE separation of proteins on a NuPAGE 4–12% gel using an ImageQuant LAS 4000 (GE Healthcare). The intensity of stained bands were quantified using ImageJ, TONSL intensity was normalized to the combined intensity of H3, H2A and H2B. Statistical analysis was performed using data from the six independent experiments using the unpaired *t*-test with equal standard deviations in prism 6.

Cell culture, transfection and drug treatment. U-2-OS (gift from J. Bartek), HeLa S3 (gift from P. Nakatami) and TIG-3 (gift from K. Hansen) cells were grown in DMEM (Gibco) containing 10% FBS (Hyclone) and 1% penicillin/streptomycin and drugs for selection. The construct for siRNA-resistant GFP–TONSL was described[2] and ARD mutations were introduced in this construct by site-directed mutagenesis. The construct for pBABE-SNAP-HA-H4 plasmid was previously described[36] and H4 tail mutations were introduced in this construct by site-directed mutagenesis. Cells inducible for GFP–TONSL wild type and ARD mutants were generated in Flp-In T-Rex U-2-OS cells (Invitrogen) by transfection of pcDNA5/FRT/TO-GFP-TONSL plasmids with Lipofectamine 2000, according to the manufacturer's protocol, and selection with hygromycin (200 μg ml^{-1}). Previously described inducible GFP–TONSL U-2-OS cells[1] were used for Fig. 1h and Fig. 2c, e. U-2-OS Flag–HA–MCM2 wild type and Y81A and Y90A cells were previously described[7]. pBABE-AsiSI–ER–HA[29] was introduced in inducible GFP–TONSL cell lines by lentiviral infection and puromycin selection. All cell lines were authenticated by western blotting and/or immunofluorescence. All cell lines used in this study tested negative for mycoplasma contamination. Expression of GFP–TONSL was induced by addition of 1 μg ml^{-1} of tetracycline for 24 h. U-2-OS and TIG3 cells were synchronized by a single thymidine block (2 mM) and released into S phase in the presence of 24 μM dCTP. For transient expression of GFP–TONSL or SNAP–HA–H4 (Fig. 1h, j), expression plasmids were introduced by transfection with Lipofectamine 2000 (Invitrogen) according to the manufacturer's protocol and cells harvested 24 h after transfection. siRNA transfection was performed with RNAiMax reagent (Invitrogen) according to the manufacturer's protocol. All siRNAs were used to a final concentration of 50 nM.

siRNA sequences (Sigma): siSET8#1: 5′-GUACGGAGCGCCAUGAAGU-3′; siSET8#2: 5′-ACUUCAUGGCGCUCCGUACUU-3′ (ref. 37); siMOF#1: 5′-GUGAUCCAGUCUCGAGUGA-3′ (ref. 12); siMOF#2: 5′-GUGAUCCAGUCUCGAGUGA-3′; siTONSL: 5′-GAGCUGGACUUAAGCAUGA-3′ (ref. 2).

Drug treatment was as follows. CPT: cells were either treated with 1 μM CPT for 3 h (Fig. 4a, i and Extended Data Fig. 8f) or 20 min (Fig. 4c), or with 50 nM CPT for 24 h (Fig. 4f). Hydroxyurea: cells were treated with 3 mM hydroxyurea for 2 h (Fig. 4b) or 3 h (Extended Data Fig. 8f).

Cell extracts and chromatin solubilisation. For detergent/salt soluble extracts, HeLa S3 and U-2-OS cells were washed with cold PBS, scraped and incubated for 15 min on ice in HS buffer supplemented with trichostatin A (TSA) and protease and phosphatase inhibitors (5 mM sodium fluoride, 10 mM β-glycerolphosphate, 0.2 mM sodium vanadate, 10 μg ml^{-1} leupeptin, 10 μg ml^{-1} pepstatin, 0.1 mM PMSF, Sigma). After centrifugation at 16,000*g* for 15 min at 4 °C, the supernatant was collected. To analyse chromatin-bound complexes, cells were washed twice in cold PBS, scraped and centrifuged at 1,500*g* for 10 min at 4 °C. The pellet was incubated on ice for 10 min in CSK buffer (10 mM PIPES pH 7, 100 mM NaCl, 300 mM sucrose, 3 mM $MgCl_2$)/0.5% Triton X-100, supplemented with TSA and protease and phosphatase inhibitors (5 mM sodium fluoride, 10 mM β-glycerolphosphate, 0.2 mM sodium vanadate, 10 μg ml^{-1} leupeptin, 10 μg ml^{-1} pepstatin, 0.1 mM PMSF, Sigma) and subsequently centrifuged at 1,500*g* for 10 min to collect soluble proteins. For DNaseI or benzonase release of chromatin material, the remaining pellet was resuspended in CSK/0.1% Triton X-100 containing DNase I (1,000 U ml^{-1}, Roche), or benzonase (2,500 U ml^{-1}, Millipore), and incubated at 30 °C for 30 min. Solubilized chromatin was then collected by centrifugation at 16,000*g* for 10 min.

Immunoprecipitation from cell extracts. Immunoprecipitation was performed with agarose magnetic GFP–Trap beads (Chromotek), anti-Flag magnetic beads (Sigma) and anti-HA magnetic beads (Life Technologies). Cell extracts were incubated with beads for 2 h at 4 °C rotating. The beads were subsequently washed five times with HS buffer and resuspended in 1× LSB before boiling and SDS–PAGE separation of proteins on a NuPAGE 4–12% gel.

Western blotting and antibodies. The following antibodies were used: TONSL (Abcam ab101898), TONSL (Sigma, HPA024679; validated in Extended Data Fig. 6d), MMS22L[1], H3 (Abcam ab1791, Abcam ab10799), GFP (Santa Cruz sc-8334, Abcam ab290), biotin (Abcam ab53494), MCM2 (BD Biosciences 610701), H2B (Abcam ab1790), H4K16ac (Millipore 07-329), H4K20me1 (Abcam ab9051), H4K20me2 (Cell Signaling 9759), 53BP1 (Santa Cruz sc-22760; Novus Biologicals NB100-904), γ-H2AX (Millipore 05-636), Cyclin B (BD Biosciences 610220), RPA70 (Abcam ab79398), SET8 (Millipore, 06-1304), MCM3 (Abcam ab 4460). Secondary antibodies conjugated with horseradish peroxidase (HRP) were from Jackson ImmunoResearch Labs. Signals were revealed by chemiluminescence substrate from Pierce (SuperSignal West Pico or SuperSignal West Femto).

FACS and analysis. For analysis of cell cycle progression, cells were fixed in 70% ethanol and stained with propidum iodide/RNase for 30 min in the dark, before analysis on a FACS Calibur machine. FACS profiles were analysed by FlowJo 10.0.8 software.

Mass spectrometry. Histones from TIG3 fibroblasts were extracted from chromatin as previously described[18]. Protein was resuspended in 50 μl of 100 mM triethylammonium bicarbonate (TEAB; Sigma), pH adjusted with 2 μl of 1.5 M Tris pH 8 and digested for 16 h at 37 °C with 3 μl of 20 ng μl^{-1} Asp-N (Wako) in 100 mM TEAB. After 15 min centrifugation at 10,000*g* at 25 °C, the supernatant was placed in a new tube and digestion was repeated for the pellet during 4 h under the conditions described earlier. The digested peptides of both digestions were merged, acidified with 10 μl of 1% TFA and purified using sequential Stagetip C18 and Carbon Toptip (Glygen). Purified peptides were evaporated, resuspended in 15 μl of 0.1% TFA. Injected material was normalized to analyse by liquid chromatography mass spectrometry (LC-MS) the histones corresponding to 9.0×10^5 cells. The LC method was used as described elsewhere[18]. The MS was performed in an Orbitrap Classic with similar settings as described previously[18] but with survey scan range at 550–690 m/z and MS2 set in scheduled and targeted data independent mode for the four-time charged ions of the four different methylation states (unmodified, mono-, di- and trimethylated H4K20) for the five different acetylated forms of S1–R23 peptide of H4 (N terminal, K5, K8, K12 and K16). Peptides were quantified using the peak area from the corresponding extracted ion chromatograms (±10 p.p.m.).

Immunofluorescence, microscopy and laser microirradiation. U-2-OS cells conditional for GFP–TONSL were grown on glass coverslips or 96-well plates and either directly fixed in 4% paraformaldehyde (PFA) for 10 min or washed in CSK, pre-extracted 5 min with cold CSK/0.5% Triton X-100 and rinsed with CSK and PBS before fixation in 4% PFA for 10 min. Coverslips were mounted on glass slides with Mowiol mounting medium (Sigma-Aldrich) containing DAPI. Fluorescence images were collected on a DeltaVision system with a ×40 or ×60 oil immersion objective. For colocalization analysis by deconvolution microscopy, z-stacks were acquired (step of 0.2 μm), deconvolved and analysed by SoftWoRX 5.0.0. Pearson coefficient correlation analysis was performed on single cells using SoftWoRX 5.0.0. Brightness and contrast were adjusted using Adobe Photoshop CS6. For high-content quantitative analysis, fluorescence images were acquired using an Olympus ScanR high-content microscope and processed on the ScanR analysis software. More than 5,000 cells per sample were analysed. Cell cycle phases were gated on DAPI and EdU intensity. Graphs were generated with TIBCO Spotfire software. For microirradiation experiments, cells grown on glass coverslips were fixed in 4% formaldehyde for 15 min, permeabilized with PBS containing 0.2% Triton X-100 for 5 min and incubated with primary antibodies diluted in DMEM for 1 h at room temperature. After staining with secondary antibodies (Alexa Fluor 488, 568 and 647; Life Technologies) for 30 min, coverslips were mounted

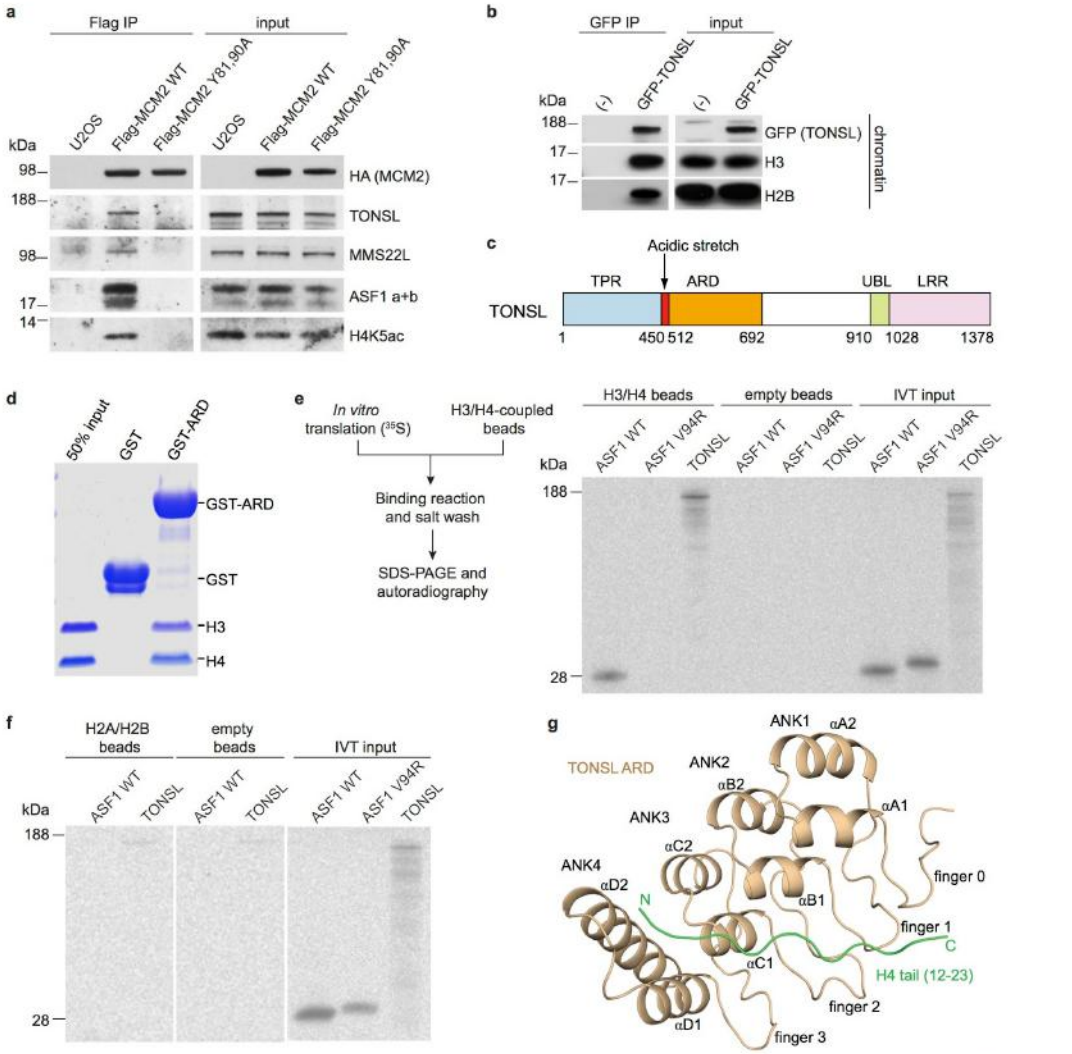
on glass slides in Vectashield mounting medium (Vector Laboratories) containing the nuclear stain DAPI. For detection of nucleotide incorporation during DNA replication, an EdU-Plus labelling kit (Life Technologies) was used according to the manufacturer's instructions. Confocal images were acquired on an LSM-780 (Carl Zeiss) mounted on a Zeiss-AxioObserver Z1 equipped with a Plan-Neofluar ×40/1.3 oil immersion objective. Image acquisition and analysis was carried out with LSM-ZEN software. Laser microirradiation of cells was performed essentially as described[38].

ChIP. GFP–TONSL wild type and N571A U-2-OS harbouring the inducible ER–HA–AsiSI endonuclease[29] were treated with 4-OHT and 10 μM DNA-PK inhibitor NU7026 (Millipore) for 4 h to increase homologous recombination[39]. Cells were cross-linked for 10 min in 1% formaldehyde and chromatin was fragmented by sonication using Bioruptor Sonicator (Diagenode). ChIP was performed as previously described[40] with the following modifications: 30 μg of chromatin was immunoprecipitated with 5 μg of anti-GFP (Abcam ab290) and rabbit-IgG. Immunoprecipitated DNA was analysed in duplicate by RT–qPCR. In all cases, γH2A.X induction was verified by immunofluorescence and a sample without 4-OHT was included as a 'no cut' control. Primer pairs for the analysis of DSB-3, DSB-I and DSB-II are described[41]. Primer sequences used for the amplification of a genomic region devoid of DSBs were as follows: noDSB-for: 5′-TGACAAGGACAGGGTCTTCC; noDSB-rev: 5′-CACCGTCCG TTGTATGTCTG. ChIP efficiency was calculated as percentage of input DNA immunoprecipitated.

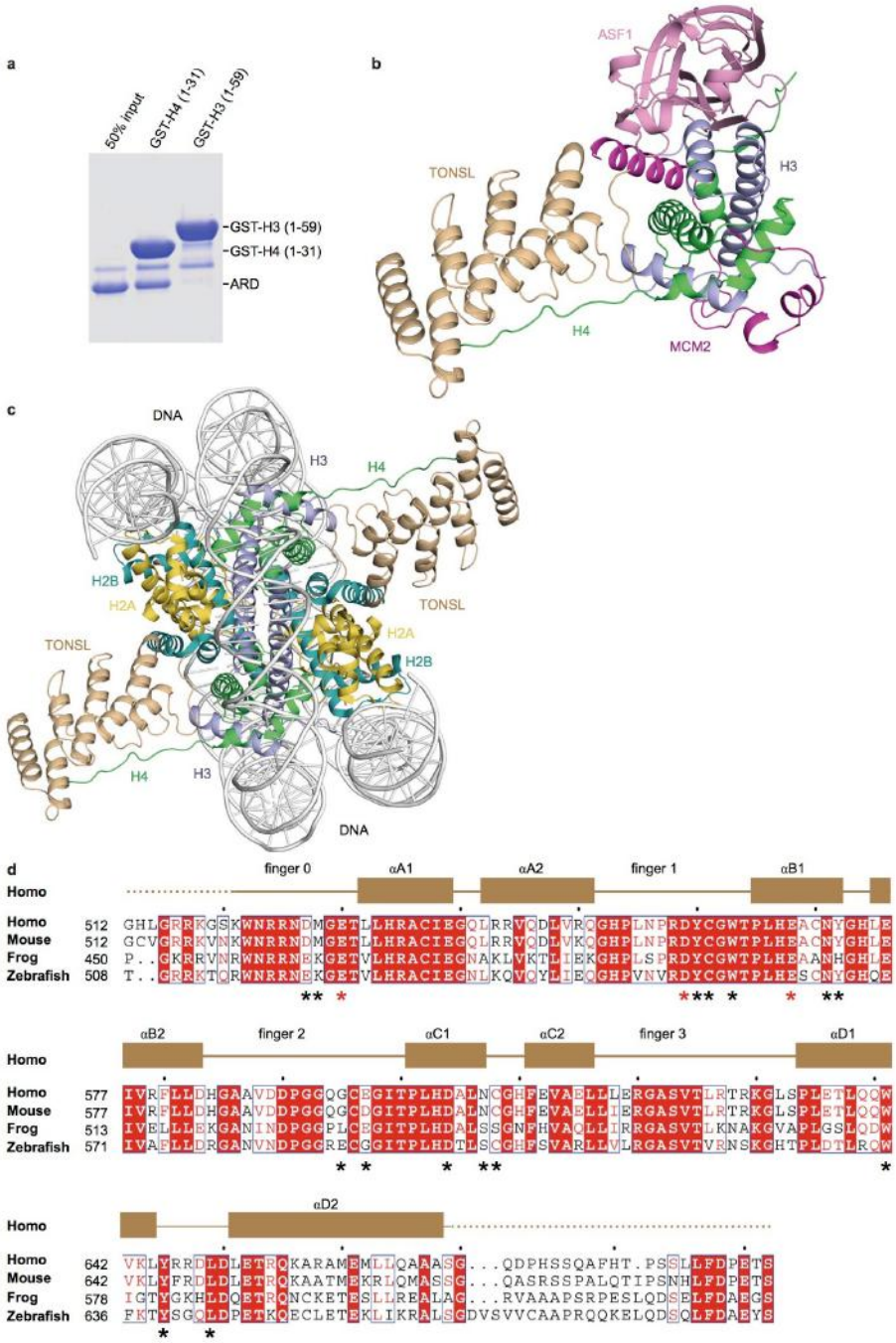
NCC. The NCC protocol[19] was adjusted for adherent U-2-OS cells. CPT (1 μM) was added 5 min before b-dUTP labelling and was included in all steps until fixation. Cells were incubated for 5 min in a hypotonic buffer (50 mM KCl, 10 mM HEPES) containing b-dUTP and resuspended into fresh cell culture medium for an additional 15 min. Cells were fixed 15 min in 1% formaldehyde, rinsed twice in PBS and collected by scraping in cold room. Nuclei were mechanically isolated in sucrose buffer (0.3 M sucrose, 10 mM HEPES-NaOH at pH 7.9, 1% Triton X-100 and 2 mM MgOAc). Chromatin was solubilized by 28 cycles 30 s on, 90 s off in sonication buffer (10 mM HEPES-NaOH at pH 7.9, 100 mM NaCl, 2 mM EDTA at pH 8, 1 mM EGTA at pH 8, 0.2% SDS, 0.1% sodium sarkosyl and 1 mM phenylmethylsulfonylfluoride) using a Bioruptor at 4 °C. Solubilized chromatin was pre-cleared using streptavidin-coated magnetic beads (MyC1 Streptavidin beads) pre-incubated with biotin. b-dUTP labelled chromatin was next purified over night at 4 °C using streptavidin-coated magnetic beads. Beads were washed five times for 2 min in wash buffer (10 mM HEPES-NaOH pH 7.9; 200 mM NaCl; 2 mM EDTA pH 8; 1 mM EGTA pH 8; 0.1% SDS; 1 mM PMSF). Total chromatin (input) and isolated nascent chromatin were boiled for 40 min on beads in LSB 1× (50 mM Tris-HCl pH 6.8, 100 mM DTT, 2% SDS, 8% glycerol, bromophenol blue) and separated by SDS–PAGE for western blotting. Pulse-SILAC-NCC (Fig. 3a) was performed as described[18].

Clonogenic assay. U-2-OS inducible for GFP–TONSL ARD wild type and mutant were transfected with siRNA, trypsinized 24 h later and seeded in technical triplicates of 1,000 or 3,000 cells in the presence or absence of tetracycline. After 24 h, cells were washed to remove tetracycline and CPT was added for 24 h as indicated. Cells were then cultured in fresh medium for 12–15 days before fixation and staining with MeOH/Crystal Violet. Colony formation efficiency was determined by manual colony counting or quantification of Crystal Violet staining by ImageJ software and normalized to non-induced control. Each data point represents a technical triplicate of 1,000 or 3,000 seeded cells within each biological replicate.

30. Cejka, P. & Kowalczykowski, S. C. The full-length *Saccharomyces cerevisiae* Sgs1 protein is a vigorous DNA helicase that preferentially unwinds holliday junctions. *J. Biol. Chem.* **285,** 8290–8301 (2010).
31. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40,** 658–674 (2007).
32. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60,** 2126–2132 (2004).
33. Adams, P. D. *et al.* PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D* **58,** 1948–1954 (2002).
34. Dyer, P. N. *et al.* Reconstitution of nucleosome core particles from recombinant histones and DNA. *Methods Enzymol.* **375,** 23–44 (2004).
35. Bartke, T. *et al.* Nucleosome-interacting proteins regulated by DNA and histone methylation. *Cell* **143,** 470–484 (2010).
36. Bodor, D. L., Valente, L. P., Mata, J. F., Black, B. E. & Jansen, L. E. Assembly in G1 phase and long-term stability are unique intrinsic features of CENP-A nucleosomes. *Mol. Biol. Cell* **24,** 923–932 (2013).
37. Jørgensen, S. *et al.* The histone methyltransferase SET8 is required for S-phase progression. *J. Cell Biol.* **179,** 1337–1345 (2007).
38. Mosbech, A., Lukas, C., Bekker-Jensen, S. & Mailand, N. The deubiquitylating enzyme USP44 counteracts the DNA double-strand break response mediated by the RNF8 and RNF168 ubiquitin ligases. *J. Biol. Chem.* **288,** 16579–16587 (2013).
39. Nakamura, K. *et al.* Regulation of homologous recombination by RNF20-dependent H2B ubiquitination. *Mol. Cell* **41,** 515–528 (2011).
40. Jakobsen, J. S. *et al.* Temporal mapping of CEBPA and CEBPB binding during liver regeneration reveals dynamic occupancy and specific regulatory codes for homeostatic and cell cycle gene batteries. *Genome Res.* **23,** 592–603 (2013).
41. Aymard, F. *et al.* Transcriptionally active chromatin recruits homologous recombination at DNA double-strand breaks. *Nature Struct. Mol. Biol.* **21,** 366–374 (2014).

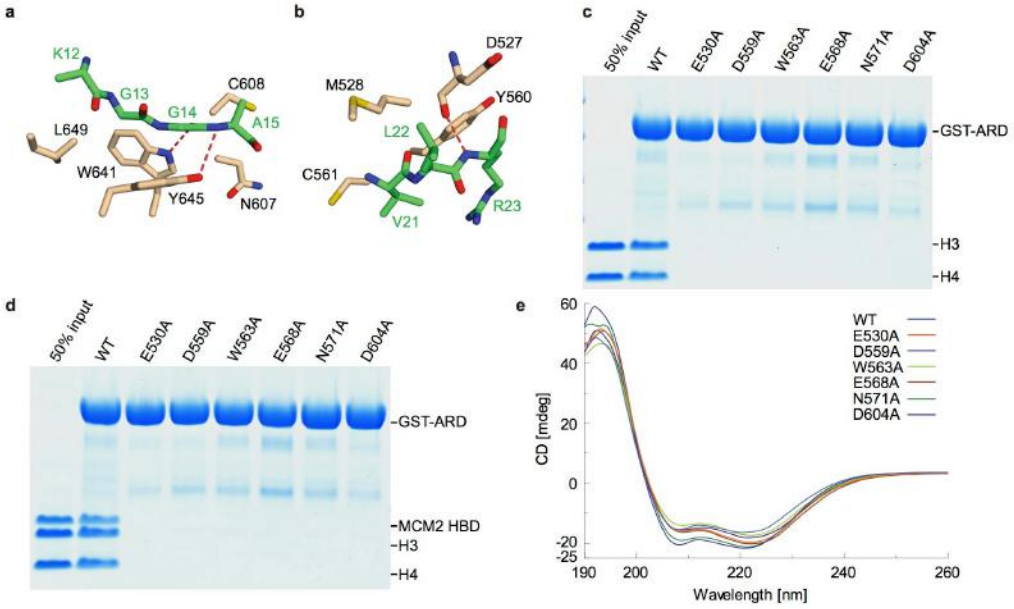


Extended Data Figure 1 | TONSL binding to histones *in vivo* and *in vitro*. **a**, Histones bridge the interaction between TONSL–MMS22L and MCM2 in cell extracts as shown by co-immunoprecipitation of Flag–HA–MCM2 wild type or histone-binding mutant (Y81A, Y90A)[7]. U-2-OS cell inducible for Flag–HA–MCM2 wild type or Y81A, Y90A[7] were induced for 24 h before immunoprecipitation with Flag antibodies (one representative experiment out of two is shown). **b**, Immunoprecipitation of GFP–TONSL from solubilized chromatin of HeLa cells transiently transfected with GFP–TONSL plasmid, showing that TONSL associates with nucleosomal histones H3 and H2B (one representative experiment out of two is shown). **c**, Domain structure of TONSL. LRR, leucine-rich repeats[1–4]; TPR, tetratricopeptide repeats; UBL, ubiquitin-like domain. **d**, Pull-down of GST–ARD with recombinant histones H3–H4 tetramers. **e**, **f**, Pull-down of *in vitro*-translated full-length TONSL with recombinant histones H3–H4 tetramers (**e**) or H2A–H2B dimers (**f**) coupled to NHS-activated sepharose beads (one representative experiment out of three (**e**) and two (**f**) is shown). ASF1a wild type and histone-binding mutant (V94R)[1] were included as controls. **g**, TONSL ARD consists of four ankyrin repeats and uses its elongated concave surface to target the H4 tail spanning residues 12 to 23.

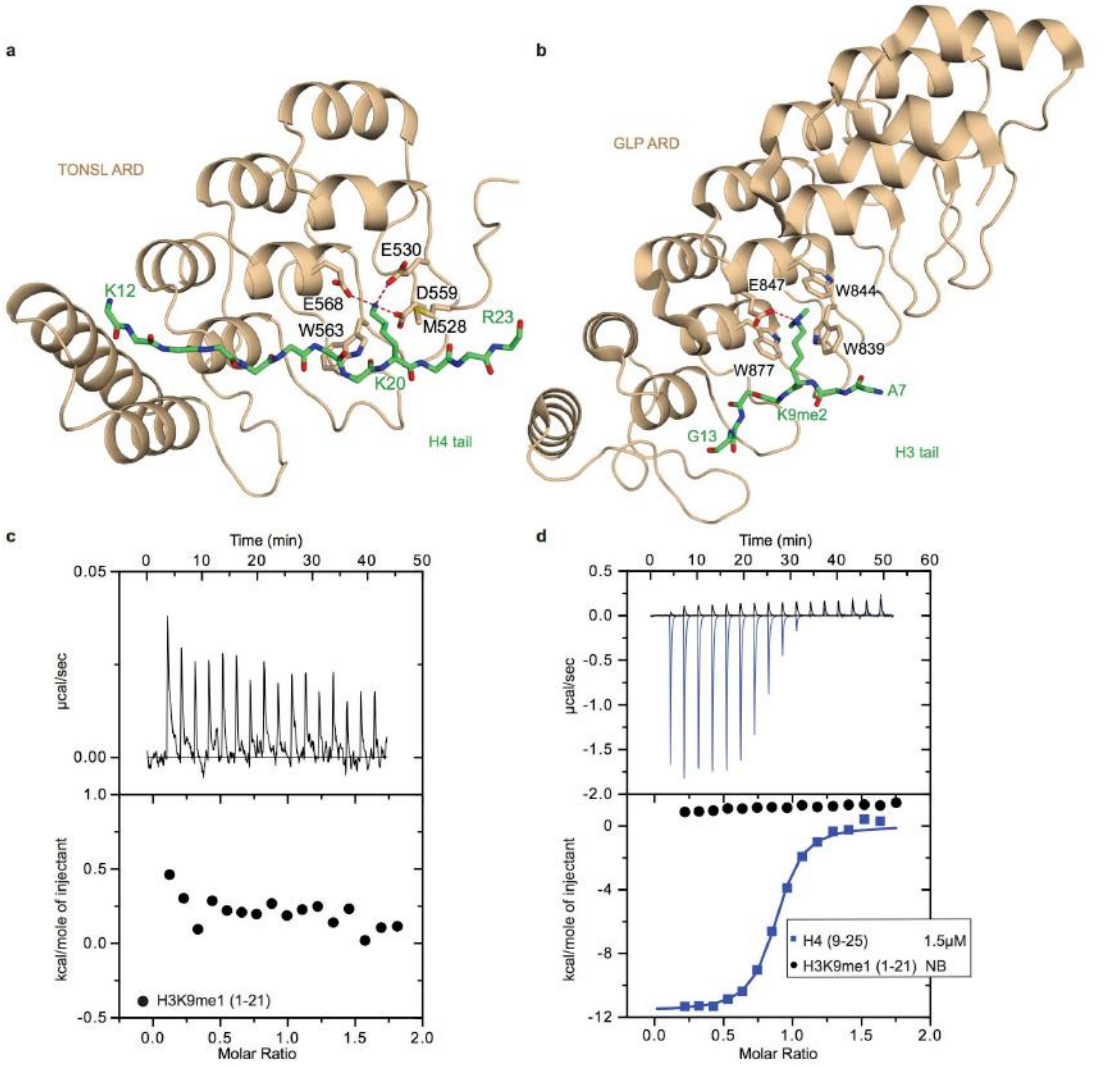


Extended Data Figure 2 | See next page for caption.

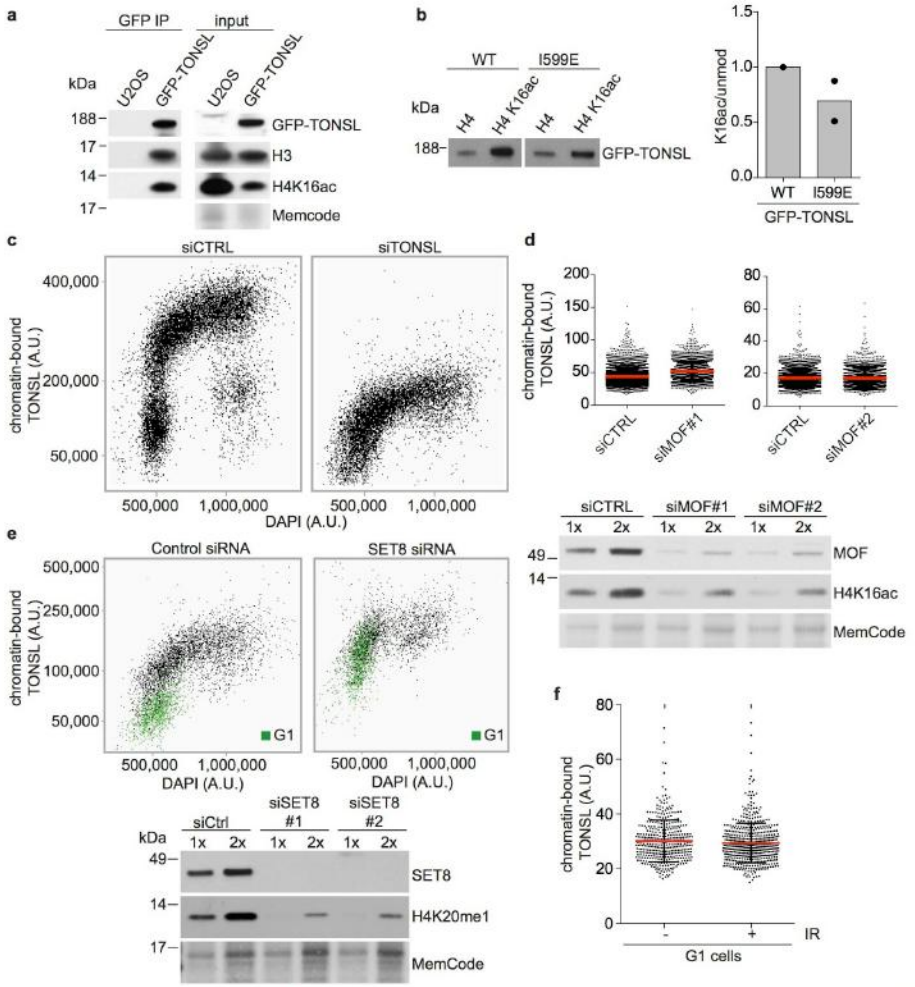
Extended Data Figure 2 | Models and sequence alignment of TONSL ARD. a, Pull-down assay of recombinant ARD with GST–H3 tail (amino acids 1–59) and GST–H4 tail (amino acids 1–31). **b**, Modelling of TONSL ARD on the co-chaperone structure of MCM2 HBD and ASF1 in complex with an H3–H4 dimer. When comparing the structure of the TONSL-ARD–MCM2-HBD–H3–H4 tetramer complex with our previous structure of the MCM2-HBD–H3–H4-dimer–ASF1 complex[7] (Protein Data Bank accession 5BNX), the common parts of both structures superimposed well with a small root mean squared deviation (r.m.s.d.) of 0.44 Å. A model of the quinary complex composed of one molecule of each protein, TONSL ARD, MCM2 HBD, ASF1, H3 and H4, was made after superposition. This model shows that TONSL ARD, MCM2 HBD and ASF1 could simultaneously bind an H3–H4 dimer without steric clash. **c**, Model of TONSL ARD on the structure of the nucleosome. The model was generated by a direct superposition of the H3–H4 tetramer in the structure of the TONSL ARD–MCM2 HBD–H3–H4 tetramer complex onto the H3–H4 tetramer in the nucleosome structure (Protein Data Bank accession 3AV2). There was no adjustment in the conformation of the model and no steric clash in the model. The MCM2 HBD molecules were omitted from the model for clarity. **d**, Alignment of TONSL ARD (512–692) sequences from *Homo sapiens*, *Mus musculus*, *Xenopus laevis* and *Danio rerio*. The secondary structures of human TONSL ARD are showed on top of the sequence alignment. Asterisks indicate the highly conserved residues that constitute the H4 tail-binding surface of TONSL ARD and the three strictly conserved acidic residues forming hydrogen bonds with the key residue H4 Lys20 are highlighted with red asterisks.



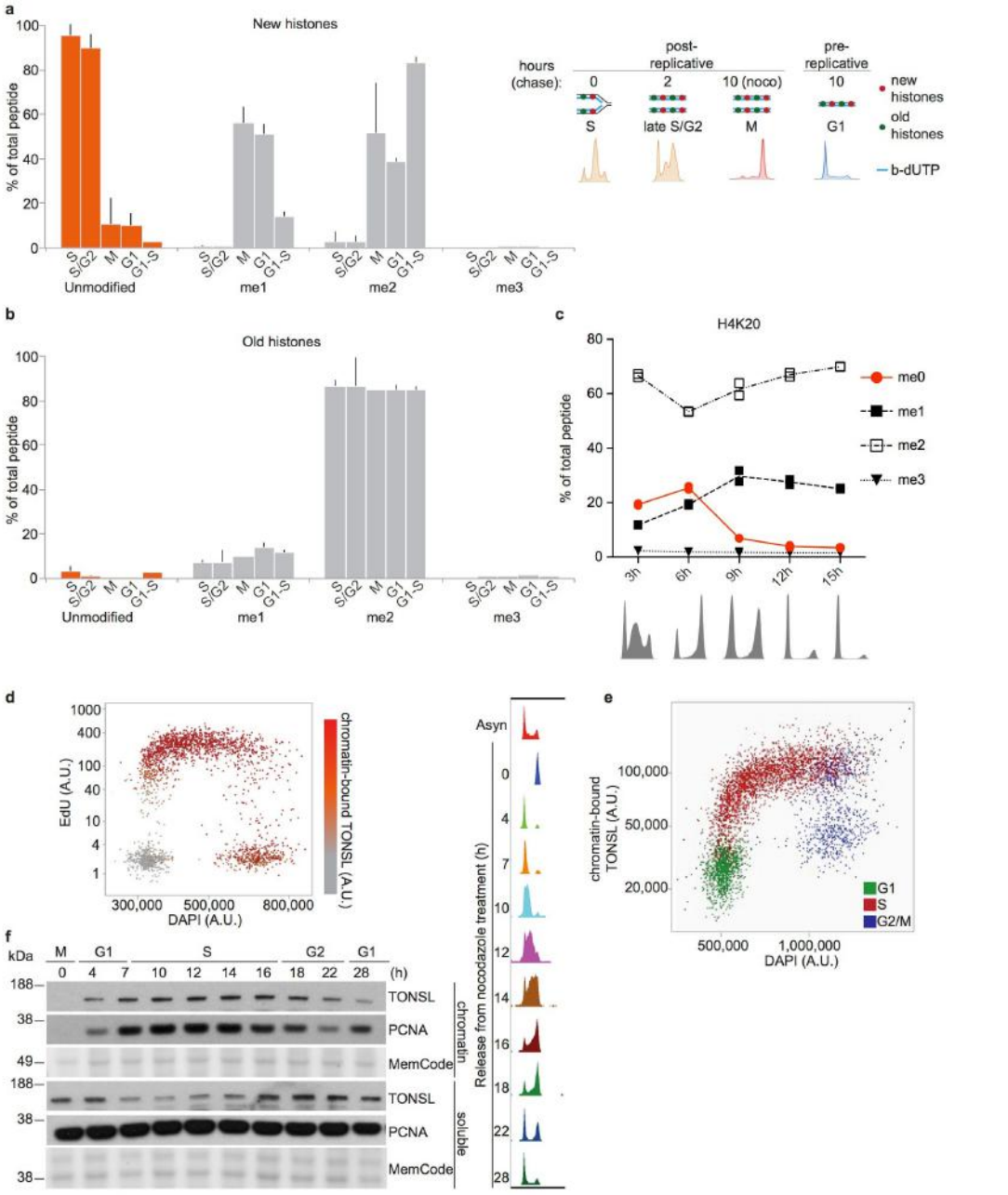
Extended Data Figure 3 | Interaction details of TONSL ARD and GST pull-downs. a, **b**, Molecular details of the interactions of TONSL ARD with H4 tail region residues 12–15 (**a**) and residues 21–23 (**b**). The Lys12-Gly13-Gly14-Ala15 segment of H4 is positioned within a narrow surface channel of the TONSL ARD scaffold. The intermolecular contacts spanning the Lys12-Gly13-Gly14-Ala15 segment of H4 include hydrophobic interactions between residues Gly13, Gly14 and Ala15 of H4 and residues Asn507, Cys508, Trp641, Tyr645 and Leu649 of ARD, as well as hydrogen bonds between the main-chain O of H4 Gly14 and Nε1 of ARD Trp641, and between the main-chain N of H4 Ala15 and Oη of ARD Tyr645 (**a**; Fig. 1c). The main-chain O of H4 Lys16 hydrogen bonds with the Nδ2 of ARD Asn571, while the side chain of H4 Lys16 forms contacts with ARD Asn607 and electrostatic interactions with the side chain of ARD Glu597 (Fig. 1c). The side chain of H4 Arg17 stacks over the side chains of ARD Tyr572 and Cys608, while its Nη1 atom forms two hydrogen bonds with main-chain O and Oδ1 of ARD Asn571 (Fig. 1c, e). The side chain of H4 H18 penetrates into a pocket lined by four strictly conserved residues (Trp563, Glu568, Asn571 and Asp604) and is positioned over His567 of ARD (Fig. 1c, f). The side chain of H4 His18 is stacked between Trp563 and Asn571 and forms hydrogen bonds to Glu568 and Asp604 of ARD (Fig. 1f). The main-chain O of H4 Arg19 forms a hydrogen bond with Nε1 of Trp563 and its side chain forms contacts with Cys561 and Gly595 of ARD (Fig. 1c). Interactions with the key residue H4 Lys20 are described in the text (Fig. 1g). The intermolecular contacts spanning the Val21-Leu22-Arg23 segment of H4 include contacts between side chains of H4 Val21 with Tyr560 and Cys561 of ARD (**b**), while H4 Leu22 interacts with Asp527 and Met528 of ARD. The main-chain N of H4 Arg23 forms a hydrogen bond with the main-chain O of Asp527 of ARD, while the side chain packs against the side chain of Tyr560 of ARD (**b**). **c**, Pull-down of recombinant histones H3–H4 with GST–TONSL ARD wild type or indicated mutants. **d**, Pull-down of pre-purified MCM2 HBD–H3–H4 tetramer complex with GST–TONSL ARD wild type or indicated mutants. **e**, Circular dichroism analysis of TONSL ARD wild type and the indicated ARD mutants.



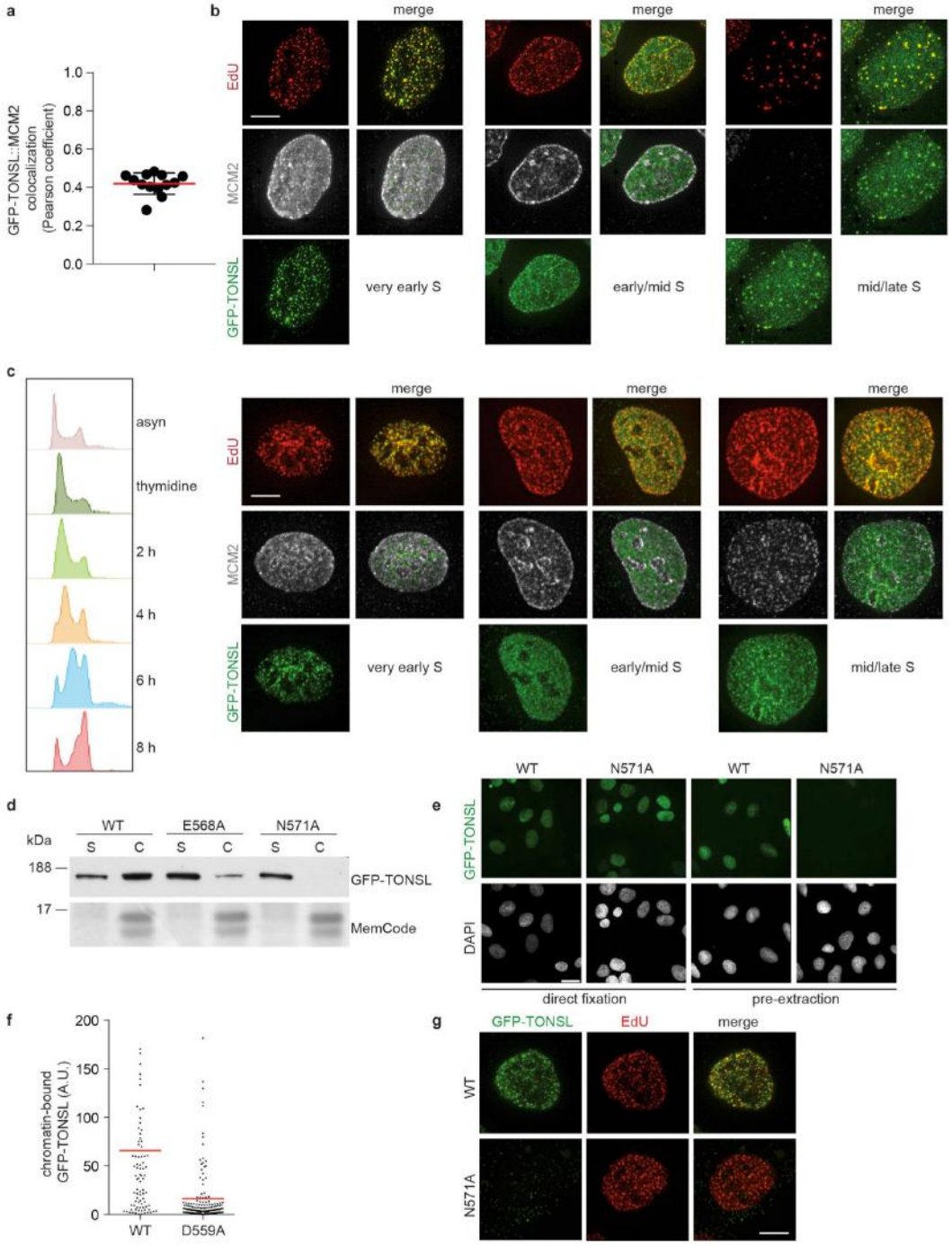
Extended Data Figure 4 | Structural comparison of the ARDs of TONSL and GLP. a, b, Representative view of the TONSL ARD with histone H4 tail (**a**; this work), and crystal structure of the GLP ARD in complex with histone H3 tail dimethylated at Lys9 (**b**; ref. 10). Both TONSL ARD and GLP ARD use the concave surface to bind their cognate target H4 tail and H3 tail, respectively. TONSL ARD recognizes H4K20me0 mainly through three strong hydrogen bonds with acidic residues Glu530, Asp559 and Glu568, while GLP ARD recognizes H3K9me2 mainly through an aromatic cage forming by residues Trp839, Trp844, Glu847 and Trp877. **c**, ITC analysis of TONSL ARD binding to H3K9me1 peptide. **d**, ITC analysis of TONSL acidic stretch and ARD (amino acids 450–692) with H3K9me1 (amino acids 1–21) and H4 (amino acids 9–25) peptides.



Extended Data Figure 5 | Effect of SET8 and MOF depletion on TONSL chromatin binding. a, Immunoprecipitation of GFP–TONSL from solubilized chromatin of GFP–TONSL U-2-OS cells (one representative experiment out of two is shown). Same exposures are shown for input and immunoprecipitation western blots of H3 and H4K16ac. **b**, TONSL ARD preference for H4K16ac could be mediated by I599 through hydrophobic association with the K16 acetyl group as I599E ARD mutation preferentially reduces binding to H4K16ac peptides as compared to the unmodified H4 tail. Left, pull-down of GFP–TONSL from cell extracts with biotinylated H4 tail peptides. Right, quantification of the western blot, GFP–TONSL binding to the H4K16ac peptide is shown relative to the unmodified peptide. Means with individual data points are shown ($n = 2$). **c**, High-content quantitative imaging of TONSL in pre-extracted U-2-OS cells. Plots show total chromatin-bound TONSL and DAPI intensities in cells treated with control or TONSL siRNA, confirming the specificity of TONSL antibody staining. Each dot represents one nucleus. **d**–**f**, Analysis of TONSL chromatin-binding in MOF-depleted (**d**), SET8-depleted (**e**) and ionizing radiation (IR)-treated cells (**f**). Chromatin-bound TONSL was quantified by high content imaging of pre-extracted U-2-OS cells stained for endogenous TONSL. Mean TONSL intensity is shown. AU, arbitrary units. **d**, **e**, Knockdown efficiency and expected effect on histone modification were confirmed by western blotting (representative of two experiments). **e**, **f**, G1 cells were defined by gating on DAPI and EdU intensity. **f**, Cells were irradiated (1.5 Gy) and analysed 1.5 h later (representative of two experiments). **d**, **f**, Error bars indicate s.d.; **d**, from left, $n = 4{,}920$, 2,341, 3,608, 2,917; **f**, $n = 382$ (−IR), 523(+IR).

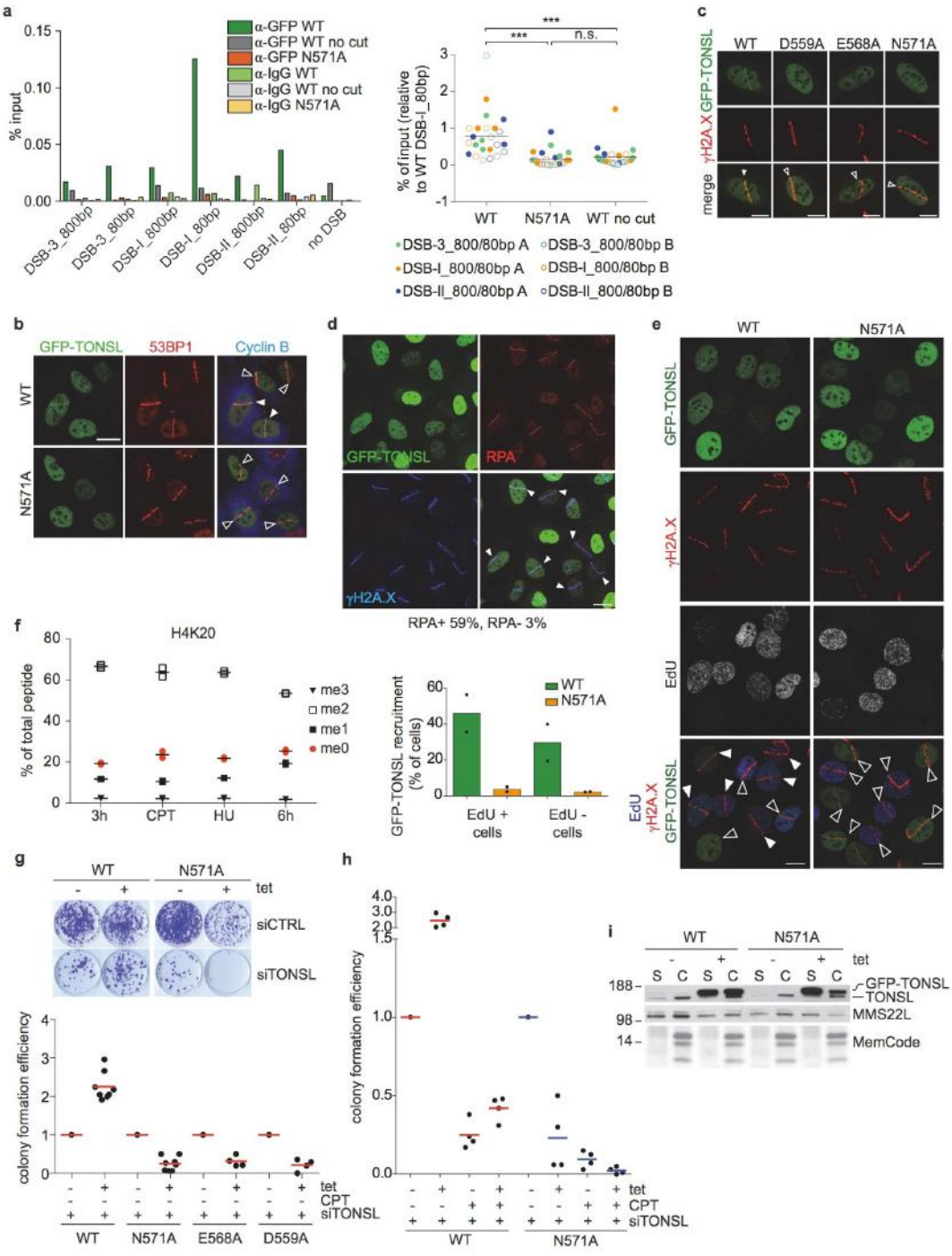


Extended Data Figure 6 | TONSL binding to chromatin during the cell cycle. a, b, H4K20 methylation levels on new and old histones analysed by NCC-pulse-SILAC (data are extracted from ref. 18). Cells grown in light SILAC medium were released into S phase in heavy medium and pulsed with b-dUTP. Chromatin was fixed, sonicated and b-dUTP-labelled fragments isolated on streptavidine beads by NCC. Histones were isolated and analysed by mass spectrometry for modifications on new (heavy) and old (light) histones. For clarity a 24 h (G1/S) chase time point is included. Error bars indicate s.d.; $n = 9$ (S), 3 (S/G2, M), 5 (G1), 3 (G1/S). Data for M (old histones) is shown as the mean of $n = 2$, as light peptides were not detected in one of the three biological replicates. **c,** H4K20 methylation levels measured by mass spectrometry in synchronized TIG3 fibroblasts. **d,** Plot of mean EdU and total DAPI intensities from TIG3 fibroblasts as in Fig. 3b, with the intensity of chromatin-bound TONSL shown in the third dimension as a colour gradient. AU, arbitrary units. Each dot represents one nucleus. Note that a population of G2 cells (EdU negative) retain TONSL on chromatin. **e,** High-content quantitative imaging of pre-extracted U-2-OS cells stained for EdU and TONSL analysed as in Fig. 3b. **f,** Analysis of TONSL chromatin binding by cellular fractionation. U-2-OS cells released from a nocodazole block were followed by fluorescence-activated cell sorting (FACS) analysis of DNA content and analysed by western blotting of soluble (CSK-Triton extracted) and chromatin (pellet) fractions (representative of two experiments).



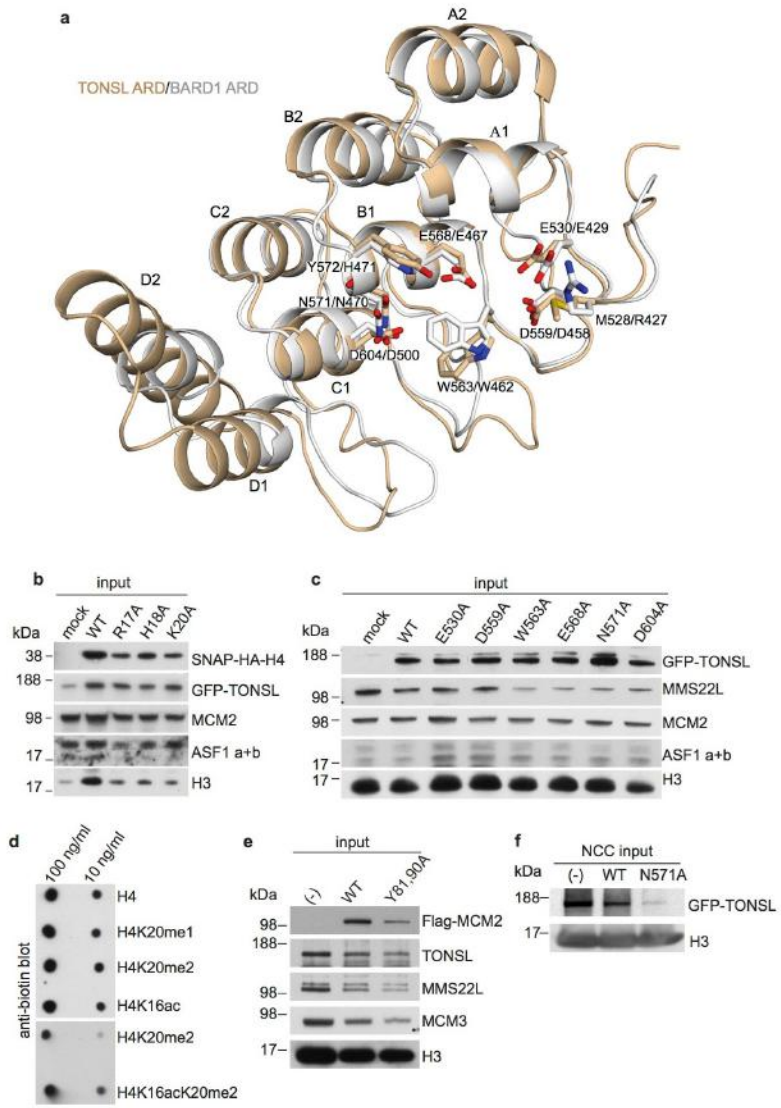
Extended Data Figure 7 | See next page for caption.

Extended Data Figure 7 | Analysis of GFP–TONSL localization. **a**, Colocalization analysis of chromatin-bound GFP–TONSL with MCM2 analysed by deconvolution microscopy and measurement of Pearson coefficient in single cells. Error bars indicate s.d., $n = 13$ from two independent experiments. Representative image, Fig. 3c. **b**, **c**, Representative images for the analysis shown in Fig. 3d. Cells were either pulsed with EdU (40 μM) for 15 min (**b**) or synchronized in G1/S and released into S phase in the continuous presence of EdU (5 μM) (**c**). Images are representative of **b**: $n = 9$ (very early), 16 (early/mid), 10 (mid/late); **c**: 9 (very early), 27 (early/mid), 36 (mid/late). Scale bar, 5 μm. **b**, EdU and MCM2 staining was used to determine the cell cycle state in asynchronous (asyn) cell populations. **c**, Progression through S phase was followed by FACS analysis of DNA content. **d**, Chromatin-binding of GFP–TONSL analysed by cellular fractionation in inducible U-2-OS cells as quantified in Fig. 3e. C, chromatin; S, soluble. **e**, **f**, Chromatin-binding analysis as in Fig. 3f. U-2-OS cells conditional for GFP–TONSL ARD wild type (WT) and mutant were directly fixed or pre-extracted to remove soluble proteins. Data are representative of three (**e**) and two (**f**) experiments, fields of cells in **e** are representative of (from left) $n = 16$, 18, 17 and 17 images. Scale bar, 20 μm. **g**, Asynchronous U-2-OS cells conditional for GFP–TONSL were pulsed with 40 μM EdU for 15 min and soluble proteins were extracted. Representative images of EdU-positive cells are shown ($n = 30$ for wild type and N571A), for the specific patterns of TONSL wild type see Fig. 3d. Scale bar, 5 μm.



Extended Data Figure 8 | See next page for caption.

Extended Data Figure 8 | TONSL–MMS22L recruitment to damaged DNA. a, Left, ChIP-qPCR analysis of GFP–TONSL recruitment to site-specific DSBs induced by AsiSI, as shown in Fig. 4d but with additional controls. Note that the colours have been changed for clarity. Mean of technical duplicates is shown. Right, dot plot illustrating the relative enrichment of GFP–TONSL wild type (WT) and N571A obtained in four independent ChIPs performed on two biologically independent chromatin preparations. Each experiment was normalized to GFP–TONSL wild-type enrichment at DSB-I_80bp. Mean is shown with two-sided Mann–Whitney test; $***P < 0.001$; not significant, $P > 0.05$; $n = 24$. Two-sided Mann–Whitney analysis of individual experiments gave similar results. **b**, U-2-OS cells conditional for GFP–TONSL were laser microirradiated. 53BP1 and cyclin B staining was used as markers of DNA damage and cells in S/G2 phase, respectively. Representative of three experiments as quantified Fig. 4e. Filled arrowheads indicate GFP–TONSL recruitment; open arrowheads indicate no recruitment. Scale bars, 10 μm **c**, U-2-OS cells transiently transfected with GFP–TONSL wild type or the indicated mutants were laser microirradiated and processed for γH2A.X immunofluorescence. Representative cells are shown ($n = 200$ cells per condition from two independent experiments). **d**, U-2-OS cells conditional for GFP–TONSL were laser microirradiated. γH2A.X and RPA staining was used as markers of DNA damage and cells undergoing resection in S/G2 phase, respectively. The percentage of GFP–TONSL cells with recruitment to RPA-positive (+) and RPA-negative (−) laser tracks is indicated. Data are representative of two independent experiments, a total of 118 cells were counted. **e**, Top, U-2-OS cells conditional for GFP–TONSL wild type and N571A were laser microirradiated. γH2A.X and EdU staining was used as markers of DNA damage and S phase cells, respectively. Bottom, quantification of GFP–TONSL cells with recruitment to laser tracks. Mean with individual data points are shown ($n = 2$, a total of 138 (wild type) and 174 (N571A) cells were counted). **f**, H4K20 methylation levels measured by mass spectrometry in synchronized TIG3 cells as in Extended Data Fig. 6c. Cell were released into S phase for 3 h and treated with hydroxyurea (HU; 3 mM) or CPT (1 μM) for 3 h or left untreated (6 h). Mean with individual data points are shown ($n = 2$). **g**, Colony formation in cells treated with control or TONSL siRNA and induced to express GFP–TONSL. As shown in Fig. 4f, but including additional mutants. Two cell concentrations in technical triplicate from two (E568A, D559A) or four (wild type, N571A) biological replicates are shown. **h**, Representation of the complementation analysis from Fig. 4f in a single panel including both CPT-treated and untreated cells. This illustrates that the toxicity of the TONSL ARD mutant is comparable to CPT treatment of cells expressing wild-type TONSL. **i**, Analysis of GFP–TONSL and MMS22L by cellular fractionation in cells inducible for GFP–TONSL ARD wild type and mutant. Representative experiment of the quantification shown in Fig. 4i.



Extended Data Figure 9 | Similarity of the ARDs in TONSL and BARD1, and protein inputs. a, Superposition of the structures of TONSL ARD and BARD1 ARD (Protein Data Bank accession 3C5R)[26]. The main residues involved in TONSL ARD interactions with the H4 tail are compared to the corresponding residues of BARD1 ARD. The two ARDs show highly similar topology and conservation of the histone-binding surface. **b**, Input material of the experiment in Fig. 1h. **c**, Input material of the experiment in Fig. 1j. **d**, Spot assay with biotinylated H4 tail (amino acids 14–33) peptides confirming equal input into pull-down reactions. **e**, Input material of the experiment in Fig. 4b. **f**, Input material of the NCC experiment in Fig. 4c. Note that because ARD mutation disrupts chromatin binding in the presence and absence of CPT (Figs 3e, f and 4a), GFP–TONSL N517 levels are low in the input chromatin. The NCC experiment in Fig. 4c supports our microscopy-based data (Fig. 4a) and further shows that there is no local accumulation of the GFP–TONSL ARD mutant at damaged forks that could have been missed in our microscopy-based quantification of total TONSL on chromatin.

Extended Data Table 1 | Data collection and refinement statistics

	TONSL ARD - MCM2 HBD - H3/H4 Tetramer Complex
Data collection	
Space group	P3 2 1
Cell dimensions	
a, *b*, *c* (Å)	139.5, 139.5, 72.9
α, β, γ (°)	90, 90, 120
Resolution (Å)	50-2.43 (2.52-2.43)*
R_{pim} (%)	3.8 (46.8)
$I/\sigma I$	23.1 (1.8)
Completeness (%)	99.8 (99.7)
Redundancy	5.5 (5.5)
Refinement	
No. reflections (total/unique)	171,308/31,146
R_{work}/ R_{free} (%)	20.1/24.6
No. atoms	
Protein	2,908
MES	12
GOL	12
Water	87
B-factors	
Protein	81.8
MES	108.5
GOL	92.6
Water	59.8
R.m.s deviations	
Bond lengths (Å)	0.009
Bond angles (°)	1.316

*Highest-resolution shell is shown in parenthesis. One crystal was used for the data.

3 CONCLUSION AND DISCUSSION

The key step in HR is the exchange of RPA assembled on ssDNA for the RAD51 recombinase, which forms a nucleoprotein filament that is required for synapsis and invasion into homologous DNA.[4,1] However, RPA-coated ssDNA possesses an inhibitory effect on RAD51 nucleoprotein filament formation. Recombination mediators, such as BRCA2 and RAD51 paralogs, including RAD51B, RAD51C, RAD51D, XRCC2, and XRCC3 are known to regulate this step of HR. [20,34,105] However, the exact mechanism and interplay in between RAD51, BRCA2 and RAD51 paralogs is not entirely clear.

Recent studies conducted by four independent groups [71-74], identified a new potential recombination mediator, MMS22L-TONSL. MMS22L-TONSL colocalized at sites of DNA damage upon laser microirradiation.[71,72] Depletion of MMS22L-TONSL did not affect resection after DNA damage. In contrast, CPT treatment of MMS22L-depleted cells led to decrease of RAD51 foci formation, showing the necessity of MMS22L-TONSL in RAD51 loading upon DNA damage.[72] Hence, MMS22L-TONSL is thought to be another recombination mediator acting alongside BRCA2 and/or RAD51 paralogs. However, whether this is true and how does the heterodimer function on a mechanistic level is not clear.

Our data brought the following conclusions:

1) MMS22L-TONSL binds long stretches of DNA, especially to RPA-coated ssDNA - Mass spectrometry analyses revealed that MMS22L-TONSL binds RPA.[71,72] Furthermore, MMS22L-TONSL colocalized with RPA foci at sites of DNA damage. [71,72] Therefore it was intriguing, whether MMS22L-TONSL binds RPA and/or RPA-coated ssDNA directly. We showed that MMS22L-TONSL binds RPA directly in solution, however the interaction was very weak (Figure S2C and D, page 82). Interestingly, once RPA is pre-bound to ssDNA, both MMS22L-TONSL and MMS22L bind RPA-coated ssDNA filament with strong affinity (Figure 2C, page 75 and Figure S2J, page 82). This suggested that RPA may serve as a recruiting platform for MMS22L-TONSL, which may serve to bring the heterodimer to sites of stalled replication forks or ssDNA gaps behind the forks.

2) MMS22L-TONSL directly binds RAD51 and RAD51-coated ssDNA - It was shown that MMS22L or TONSL depletion affects RAD51 foci formation upon treatment with CPT.[71,72] However the underlying mechanism was not clear. Specifically, it was not clear if MMS22L-TONSL interacts with RAD51. Here, we showed that MMS22L-TONSL and MMS22L alone bind RAD51 directly (5F and G, page 78; S5K, page 85). Furthermore, our additional data provide evidence that MMS22L-TONSL also binds RAD51-coated ssDNA filament (see section 2.2.3, Figure 3, page 95). This suggests that the effect of MMS22L-TONSL might be rather direct and may depend on the physical interaction between MMS22L-TONSL and RAD51.

3) MMS22L-TONSL lowers RAD51 binding affinity to dsDNA and promotes RAD51-dependent strand exchange activity - Downregulation of MMS22L and TONSL led to approximately 60% reduction in GFP-positive cells in DR-GFP reporter assays.[72] This suggested that MMS22L-TONSL plays an important function in homologous recombination. Together with the RAD51 foci data, this further supports the hypothesis that MMS22L-TONSL is a factor, which plays a role in the loading of RAD51 onto RPA-coated ssDNA filament.[72] However, *in vivo* data cannot substantiate/uncover the exact molecular mechanism how MMS22L-TONSL functions. Recombination mediators can promote RAD51 loading in multiple ways, for example - (i) by overcoming inhibitory effect of RPA, (ii) stabilizing the RAD51-ssDNA nucleoprotein filament, and (iii) by inhibiting RAD51's binding affinity to dsDNA, which is inhibitory for the strand exchange reaction.[20,34,54] Here we showed that MMS22L-TONSL lowers the binding affinity of RAD51 to dsDNA (Figures 5C, D, H, K on page 78; S5A, B, C, H on page 85). In reactions where both double and single-stranded DNA are present, MMS22L-TONSL facilitates the loading of RAD51 on ssDNA. These data altogether point out that MMS22L-TONSL is a recombination mediator in HR. It may specifically act at stalled or collapsed DNA replication forks to inhibits RAD51's binding affinity towards dsDNA (Figure 7A and B, page 80). This leads to the stimulation of DNA strand exchange in reconstituted assays. However, MMS22L-TONSL cannot overcome the inhibitory effect of RPA (Figure S5N, page 85), therefore we envision that it acts synergistically with other recombination mediators. The exact interplay and possible synergistic effect/cooperation of MMS22L-TONSL with other mediators such as RAD51 paralogs is not yet clear, and

will represent a challenge for future research. Furthermore, the effect of MMS22L-TONSL on recombination specifically in the context of stalled replication forks will need to be established.

4 REFERENCES

1 Krejci, L., Altmannova, V., Spirek, M. & Zhao, X. Homologous recombination and its regulation. *Nucleic Acids Res* **40**, 5795-5818, doi:10.1093/nar/gks270 (2012).

2 San Filippo, J., Sung, P. & Klein, H. Mechanism of eukaryotic homologous recombination. *Annu Rev Biochem* **77**, 229-257, doi:10.1146/annurev.biochem.77.061306.125255 (2008).

3 Sturzenegger, A. *et al.* DNA2 cooperates with the WRN and BLM RecQ helicases to mediate long-range DNA end resection in human cells. *J Biol Chem* **289**, 27314-27326, doi:10.1074/jbc.M114.578823 (2014).

4 Heyer, W. D., Ehmsen, K. T. & Liu, J. Regulation of homologous recombination in eukaryotes. *Annu Rev Genet* **44**, 113-139, doi:10.1146/annurev-genet-051710-150955 (2010).

5 Alabert, C. & Groth, A. Chromatin replication and epigenome maintenance. *Nat Rev Mol Cell Bio* **13**, 153-167, doi:10.1038/nrm3288 (2012).

6 Nimonkar, A. V. *et al.* BLM-DNA2-RPA-MRN and EXO1-BLM-RPA-MRN constitute two DNA end resection machineries for human DNA break repair. *Gene Dev* **25**, 350-362, doi:10.1101/gad.2003811 (2011).

7 Fan, J. & Pavletich, N. P. Structure and conformational change of a replication protein A heterotrimer bound to ssDNA. *Gene Dev* **26**, 2337-2347, doi:10.1101/gad.194787.112 (2012).

8 Sung, P. & Klein, H. Mechanism of homologous recombination: mediators and helicases take on regulatory functions. *Nat Rev Mol Cell Biol* **7**, 739-750, doi:10. 1038/nrm2008 (2006).

9 Li, X. & Heyer, W. D. Homologous recombination in DNA repair and DNA damage tolerance. *Cell Res* **18**, 99-113, doi:10.1038/cr.2008.1 (2008).

10 Matos, J. & West, S. C. Holliday junction resolution: regulation in space and time. *DNA Repair (Amst)* **19**, 176-181, doi:10.1016/j.dnarep.2014.03.013 (2014).

11 Sarbajna, S., Davies, D. & West, S. C. Roles of SLX1-SLX4, MUS81-EME1, and GEN1 in avoiding genome instability and mitotic catastrophe. *Gene Dev* **28**, 1124-1136, doi:10.1101/gad.238303.114 (2014).

12 Petukhova, G. V. *et al.* The Hop2 and Mnd1 proteins act in concert with Rad51 and Dmc1 in meiotic recombination. *Nat Struct Mol Biol* **12**, 449-453, doi:10.1038/nsmb923 (2005).

13 Sauvageau, S. *et al.* Fission yeast rad51 and dmc1, two efficient DNA recombinases forming helical nucleoprotein filaments. *Mol Cell Biol* **25**, 4377-4387, doi:10.1128/MCB.25.11.4377-4387.2005 (2005).

14 Shinohara, A. & Shinohara, M. Roles of RecA homologues Rad51 and Dmc1 during meiotic recombination. *Cytogenet Genome Res* **107**, 201-207, doi:10.1159/000080598 (2004).

15 Hiom, K. DNA repair: common approaches to fixing double-strand breaks. *Curr Biol* **19**, R523-525, doi:10.1016/j.cub.2009.06.009 (2009).

16 Wigley, D. B. Bacterial DNA repair: recent insights into the mechanism of RecBCD, AddAB and AdnAB. *Nat Rev Microbiol* **11**, 9-13, doi:10.1038/nrmicro2917 (2013).

17 Dillingham, M. S. & Kowalczykowski, S. C. RecBCD enzyme and the repair of double-stranded DNA breaks. *Microbiol Mol Biol Rev* **72**, 642-671, Table of Contents, doi:10.1128/MMBR.00020-08 (2008).
18 Morimatsu, K. & Kowalczykowski, S. C. RecFOR proteins load RecA protein onto gapped DNA to accelerate DNA strand exchange: A universal step of recombinational repair. *Mol Cell* **11**, 1337-1347, doi:Doi 10.1016/S1097-2765(03)00188-6 (2003).
19 Handa, N., Morimatsu, K., Lovett, S. T. & Kowalczykowski, S. C. Reconstitution of initial steps of dsDNA break repair by the RecF pathway of E. coli. *Gene Dev* **23**, 1234-1245, doi:10.1101/gad.1780709 (2009).
20 Jensen, R. B., Carreira, A. & Kowalczykowski, S. C. Purified human BRCA2 stimulates RAD51-mediated recombination. *Nature* **467**, 678-683, doi:10.1038/nature09399 (2010).
21 Liu, J. *et al.* Rad51 paralogues Rad55-Rad57 balance the antirecombinase Srs2 in Rad51 filament formation. *Nature* **479**, 245-248, doi:10.1038/nature10522 (2011).
22 Sugiyama, T., New, J. H. & Kowalczykowski, S. C. DNA annealing by RAD52 protein is stimulated by specific interaction with the complex of replication protein A and single-stranded DNA. *Proc Natl Acad Sci U S A* **95**, 6049-6054 (1998).
23 Rijkers, T. *et al.* Targeted inactivation of mouse RAD52 reduces homologous recombination but not resistance to ionizing radiation. *Mol Cell Biol* **18**, 6423-6429 (1998).
24 Fujimori, A. *et al.* Rad52 partially substitutes for the Rad51 paralog XRCC3 in maintaining chromosomal integrity in vertebrate cells. *The EMBO journal* **20**, 5513-5520, doi:10.1093/emboj/20.19.5513 (2001).
25 Feng, Z. *et al.* Rad52 inactivation is synthetically lethal with BRCA2 deficiency. *Proc Natl Acad Sci U S A* **108**, 686-691, doi:10.1073/pnas.1010959107 (2011).
26 Meindl, A. *et al.* Germline mutations in breast and ovarian cancer pedigrees establish RAD51C as a human cancer susceptibility gene. *Nat Genet* **42**, 410-414, doi:10.1038/ng.569 (2010).
27 Sy, S. M., Huen, M. S. & Chen, J. PALB2 is an integral component of the BRCA complex required for homologous recombination repair. *Proc Natl Acad Sci U S A* **106**, 7155-7160, doi:10.1073/pnas.0811159106 (2009).
28 Zhang, F. *et al.* PALB2 links BRCA1 and BRCA2 in the DNA-damage response. *Curr Biol* **19**, 524-529, doi:10.1016/j.cub.2009.02.018 (2009).
29 Moynahan, M. E. The cancer connection: BRCA1 and BRCA2 tumor suppression in mice and humans. *Oncogene* **21**, 8994-9007, doi:10.1038/sj.onc.1206177 (2002).
30 Xia, B. *et al.* Control of BRCA2 cellular and clinical functions by a nuclear partner, PALB2. *Mol Cell* **22**, 719-729, doi:10.1016/j.molcel.2006.05.022 (2006).
31 Tarsounas, M., Davies, D. & West, S. C. BRCA2-dependent and independent formation of RAD51 nuclear foci. *Oncogene* **22**, 1115-1123, doi:10.1038/sj.onc.1206263 (2003).
32 Lu, H. *et al.* The BRCA2-interacting protein BCCIP functions in RAD51 and BRCA2 focus formation and homologous recombinational repair. *Mol Cell Biol* **25**, 1949-1957, doi:10.1128/MCB.25.5.1949-1957.2005 (2005).

33 Schlacher, K. *et al.* Double-strand break repair-independent role for BRCA2 in blocking stalled replication fork degradation by MRE11. *Cell* **145**, 529-542, doi:10.1016/j.cell.2011.03.041 (2011).
34 Thorslund, T. *et al.* The breast cancer tumor suppressor BRCA2 promotes the specific targeting of RAD51 to single-stranded DNA. *Nat Struct Mol Biol* **17**, 1263-1265, doi:10.1038/nsmb.1905 (2010).
35 Roy, R., Chun, J. & Powell, S. N. BRCA1 and BRCA2: different roles in a common pathway of genome protection. *Nature reviews. Cancer* **12**, 68-78, doi:10.1038/nrc3181 (2012).
36 Lord, C. J. & Ashworth, A. RAD51, BRCA2 and DNA repair: a partial resolution. *Nat Struct Mol Biol* **14**, 461-462, doi:10.1038/nsmb0607-461 (2007).
37 Chatterjee, G., Jimenez-Sainz, J., Presti, T., Nguyen, T. & Jensen, R. B. Distinct binding of BRCA2 BRC repeats to RAD51 generates differential DNA damage sensitivity. *Nucleic Acids Res*, doi:10.1093/nar/gkw242 (2016).
38 Shahid, T. *et al.* Structure and mechanism of action of the BRCA2 breast cancer tumor suppressor. *Nat Struct Mol Biol* **21**, 962-968, doi:10.1038/nsmb.2899 (2014).
39 Milner, J., Ponder, B., Hughes-Davies, L., Seltmann, M. & Kouzarides, T. Transcriptional activation functions in BRCA2. *Nature* **386**, 772-773, doi:10.1038/386772a0 (1997).
40 Wong, J. M. S., Ionescu, D. & Ingles, C. J. Interaction between BRCA2 and replication protein A is compromised by a cancer-predisposing mutation in BRCA2. *Oncogene* **22**, 28-33, doi:10.1038/sj.onc.1206071 (2003).
41 Bugreev, D. V. & Mazin, A. V. Ca2+ activates human homologous recombination protein Rad51 by modulating its ATPase activity. *Proc Natl Acad Sci U S A* **101**, 9988-9993, doi:10.1073/pnas.0402105101 (2004).
42 Liu, J., Doty, T., Gibson, B. & Heyer, W. D. Human BRCA2 protein promotes RAD51 filament formation on RPA-covered single-stranded DNA. *Nat Struct Mol Biol* **17**, 1260-1262, doi:10.1038/nsmb.1904 (2010).
43 Gudmundsdottir, K. & Ashworth, A. The roles of BRCA1 and BRCA2 and associated proteins in the maintenance of genomic stability. *Oncogene* **25**, 5864-5874, doi:10.1038/sj.onc.1209874 (2006).
44 Wold, M. S. Replication protein A: a heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism. *Annu Rev Biochem* **66**, 61-92, doi:10.1146/annurev.biochem.66.1.61 (1997).
45 Pellegrini, L. *et al.* Insights into DNA recombination from the structure of a RAD51-BRCA2 complex. *Nature* **420**, 287-293, doi:10.1038/nature01230 (2002).
46 Carreira, A. *et al.* The BRC repeats of BRCA2 modulate the DNA-binding selectivity of RAD51. *Cell* **136**, 1032-1043, doi:10.1016/j.cell.2009.02.019 (2009).
47 Carreira, A. & Kowalczykowski, S. C. Two classes of BRC repeats in BRCA2 promote RAD51 nucleoprotein filament function by distinct mechanisms. *Proc Natl Acad Sci U S A* **108**, 10448-10453, doi:10.1073/pnas.1106971108 (2011).
48 Sung, P. Yeast Rad55 and Rad57 proteins form a heterodimer that functions with replication protein a to promote DNA strand exchange by Rad51 recombinase. *Gene Dev* **11**, 1111-1121, doi:Doi 10.1101/Gad.11.9.1111 (1997).

49 Gaines, W. A. *et al.* Promotion of presynaptic filament assembly by the ensemble of S. cerevisiae Rad51 paralogues with Rad52. *Nat Commun* **6**, 7834, doi:10.1038/ncomms8834 (2015).
50 Thacker, J. The RAD51 gene family, genetic instability and cancer. *Cancer Lett* **219**, 125-135, doi:10.1016/j.canlet.2004.08.018 (2005).
51 Takata, M. *et al.* Chromosome instability and defective recombinational repair in knockout mutants of the five Rad51 paralogs. *Mol Cell Biol* **21**, 2858-2866, doi:10.1128/MCB.21.8.2858-2866.2001 (2001).
52 Suwaki, N., Klare, K. & Tarsounas, M. RAD51 paralogs: roles in DNA damage signalling, recombinational repair and tumorigenesis. *Semin Cell Dev Biol* **22**, 898-905, doi:10.1016/j.semcdb.2011.07.019 (2011).
53 Thompson, L. H. & Schild, D. Homologous recombinational repair of DNA ensures mammalian chromosome stability. *Mutat Res* **477**, 131-153 (2001).
54 Sigurdsson, S. *et al.* Mediator function of the human Rad51B-Rad51C complex in Rad51/RPA-catalyzed DNA strand exchange. *Gene Dev* **15**, 3308-3318, doi:10.1101/gad.935501 (2001).
55 Sung, P., Trujillo, K. M. & Van Komen, S. Recombination factors of Saccharomyces cerevisiae. *Mutat Res* **451**, 257-275 (2000).
56 Kurumizaka, H. *et al.* Homologous pairing and ring and filament structure formation activities of the human Xrcc2*Rad51D complex. *J Biol Chem* **277**, 14315-14320, doi:10.1074/jbc.M105719200 (2002).
57 Chun, J., Buechelmaier, E. S. & Powell, S. N. Rad51 paralog complexes BCDX2 and CX3 act at different stages in the BRCA1-BRCA2-dependent homologous recombination pathway. *Mol Cell Biol* **33**, 387-395, doi:10.1128/MCB.00465-12 (2013).
58 Masson, J. Y. *et al.* Identification and purification of two distinct complexes containing the five RAD51 paralogs. *Gene Dev* **15**, 3296-3307, doi:10.1101/gad.947001 (2001).
59 Braybrooke, J. P., Spink, K. G., Thacker, J. & Hickson, I. D. The RAD51 family member, RAD51L3, is a DNA-stimulated ATPase that forms a complex with XRCC2. *J Biol Chem* **275**, 29100-29106, doi:10.1074/jbc.M002075200 (2000).
60 Taylor, M. R. *et al.* Rad51 Paralogs Remodel Pre-synaptic Rad51 Filaments to Stimulate Homologous Recombination. *Cell* **162**, 271-286, doi:10.1016/j.cell.2015.06.015 (2015).
61 Ceballos, S. J. & Heyer, W. D. Functions of the Snf2/Swi2 family Rad54 motor protein in homologous recombination. *Biochim Biophys Acta* **1809**, 509-523, doi:10.1016/j.bbagrm.2011.06.006 (2011).
62 Mazin, A. V., Mazina, O. M., Bugreev, D. V. & Rossi, M. J. Rad54, the motor of homologous recombination. *DNA Repair (Amst)* **9**, 286-302, doi:10.1016/j.dnarep.2009.12.006 (2010).
63 Wesoly, J. *et al.* Differential contributions of mammalian Rad54 paralogs to recombination, DNA damage repair, and meiosis. *Mol Cell Biol* **26**, 976-989, doi:10.1128/MCB.26.3.976-989.2006 (2006).
64 Sarai, N. *et al.* Stimulation of Dmc1-mediated DNA strand exchange by the human Rad54B protein. *Nucleic Acids Res* **34**, 4429-4437, doi:10.1093/nar/gkl562 (2006).
65 Zhao, W. & Sung, P. Significance of ligand interactions involving Hop2-Mnd1 and the RAD51 and DMC1 recombinases in homologous DNA repair and XX

ovarian dysgenesis. *Nucleic Acids Res* **43**, 4055-4066, doi:10.1093/nar/gkv259 (2015).
66 Prakash, L. & Prakash, S. Isolation and characterization of MMS-sensitive mutants of Saccharomyces cerevisiae. *Genetics* **86**, 33-55 (1977).
67 Collins, S. R. *et al.* Functional dissection of protein complexes involved in yeast chromosome biology using a genetic interaction map. *Nature* **446**, 806-810, doi:10.1038/nature05649 (2007).
68 Zaidi, I. W. *et al.* Rtt101 and Mms1 in budding yeast form a CUL4(DDB1)-like ubiquitin ligase that promotes replication through damaged DNA. *EMBO Rep* **9**, 1034-1040, doi:10.1038/embor.2008.155 (2008).
69 Duro, E., Vaisica, J. A., Brown, G. W. & Rouse, J. Budding yeast Mms22 and Mms1 regulate homologous recombination induced by replisome blockage. *DNA Repair (Amst)* **7**, 811-818, doi:10.1016/j.dnarep.2008.01.007 (2008).
70 Ray Chaudhuri, A. *et al.* Topoisomerase I poisoning results in PARP-mediated replication fork reversal. *Nat Struct Mol Biol* **19**, 417-423, doi:10.1038/nsmb.2258 (2012).
71 Duro, E. *et al.* Identification of the MMS22L-TONSL complex that promotes homologous recombination. *Mol Cell* **40**, 632-644, doi:10.1016/j.molcel.2010.10.023 (2010).
72 O'Donnell, L. *et al.* The MMS22L-TONSL complex mediates recovery from replication stress and homologous recombination. *Mol Cell* **40**, 619-631, doi:10.1016/j.molcel.2010.10.024 (2010).
73 O'Connell, B. C. *et al.* A genome-wide camptothecin sensitivity screen identifies a mammalian MMS22L-NFKBIL2 complex required for genomic stability. *Mol Cell* **40**, 645-657, doi:10.1016/j.molcel.2010.10.022 (2010).
74 Piwko, W. *et al.* RNAi-based screening identifies the Mms22L-Nfkbil2 complex as a novel regulator of DNA replication in human cells. *The EMBO journal* **29**, 4210-4222, doi:10.1038/emboj.2010.304 (2010).
75 Blatch, G. L. & Lassle, M. The tetratricopeptide repeat: a structural motif mediating protein-protein interactions. *Bioessays* **21**, 932-939, doi:10.1002/(SICI)1521-1878(199911)21:11<932::AID-BIES5>3.0.CO;2-N (1999).
76 Kobe, B. & Kajava, A. V. The leucine-rich repeat as a protein recognition motif. *Curr Opin Struct Biol* **11**, 725-732 (2001).
77 Moscat, J., Diaz-Meco, M. T., Albert, A. & Campuzano, S. Cell signaling and function organized by PB1 domain interactions. *Mol Cell* **23**, 631-640, doi:10.1016/j.molcel.2006.08.002 (2006).
78 Campos, E. I. *et al.* Analysis of the Histone H3.1 Interactome: A Suitable Chaperone for the Right Event. *Mol Cell* **60**, 697-709, doi:10.1016/j.molcel.2015.08.005 (2015).
79 Mazouzi, A., Velimezi, G. & Loizou, J. I. DNA replication stress: causes, resolution and disease. *Exp Cell Res* **329**, 85-93, doi:10.1016/j.yexcr.2014.09.030 (2014).
80 Marechal, A. & Zou, L. DNA damage sensing by the ATM and ATR kinases. *Cold Spring Harb Perspect Biol* **5**, doi:10.1101/cshperspect.a012716 (2013).
81 Lobrich, M. & Jeggo, P. A. The impact of a negligent G2/M checkpoint on genomic instability and cancer induction. *Nature reviews. Cancer* **7**, 861-869, doi:10.1038/nrc2248 (2007).

82 Walker, M., Black, E. J., Oehler, V., Gillespie, D. A. & Scott, M. T. Chk1 C-terminal regulatory phosphorylation mediates checkpoint activation by de-repression of Chk1 catalytic activity. *Oncogene* **28**, 2314-2323, doi:10.1038/onc.2009.102 (2009).
83 Moynahan, M. E. & Jasin, M. Mitotic homologous recombination maintains genomic stability and suppresses tumorigenesis. *Nat Rev Mol Cell Biol* **11**, 196-207, doi:10.1038/nrm2851 (2010).
84 Piwko, W., Buser, R. & Peter, M. Rescuing stalled replication forks: MMS22L-TONSL, a novel complex for DNA replication fork repair in human cells. *Cell Cycle* **10**, 1703-1705, doi:10.4161/cc.10.11.15557 (2011).
85 Hanahan, D. & Weinberg, R. A. Hallmarks of Cancer: The Next Generation. *Cell* **144**, 646-674, doi:10.1016/j.cell.2011.02.013 (2011).
86 Turner, N., Tutt, A. & Ashworth, A. Hallmarks of 'BRCAness' in sporadic cancers. *Nature reviews. Cancer* **4**, 814-819, doi:10.1038/nrc1457 (2004).
87 Huang, F. *et al.* Targeting BRCA1- and BRCA2-deficient cells with RAD52 small molecule inhibitors. *Nucleic Acids Res* **44**, 4189-4199, doi:10.1093/nar/gkw087 (2016).
88 Greer, J. B. & Whitcomb, D. C. Role of BRCA1 and BRCA2 mutations in pancreatic cancer. *Gut* **56**, 601-605, doi:10.1136/gut.2006.101220 (2007).
89 Prakash, R., Zhang, Y., Feng, W. & Jasin, M. Homologous recombination and human health: the roles of BRCA1, BRCA2, and associated proteins. *Cold Spring Harb Perspect Biol* **7**, a016600, doi:10.1101/cshperspect.a016600 (2015).
90 Tischkowitz, M. & Xia, B. PALB2/FANCN: recombining cancer and Fanconi anemia. *Cancer Res* **70**, 7353-7359, doi:10.1158/0008-5472.CAN-10-1012 (2010).
91 Antoniou, A. C. *et al.* Breast-cancer risk in families with mutations in PALB2. *N Engl J Med* **371**, 497-506, doi:10.1056/NEJMoa1400382 (2014).
92 Sawyer, S. L. *et al.* Biallelic mutations in BRCA1 cause a new Fanconi anemia subtype. *Cancer Discov* **5**, 135-142, doi:10.1158/2159-8290.CD-14-1156 (2015).
93 Meyer, S. *et al.* Fanconi anaemia, BRCA2 mutations and childhood cancer: a developmental perspective from clinical and epidemiological observations with implications for genetic counselling. *J Med Genet* **51**, 71-75, doi:10.1136/jmedgenet-2013-101642 (2014).
94 Han, X., Saito, H., Miki, Y. & Nakanishi, A. A CRM1-mediated nuclear export signal governs cytoplasmic localization of BRCA2 and is essential for centrosomal localization of BRCA2. *Oncogene* **27**, 2969-2977, doi:10.1038/sj.onc.1210968 (2008).
95 Jeyasekharan, A. D. *et al.* A cancer-associated BRCA2 mutation reveals masked nuclear export signals controlling localization. *Nat Struct Mol Biol* **20**, 1191-1198, doi:10.1038/nsmb.2666 (2013).
96 Goldgar, D. E. *et al.* Integrated evaluation of DNA sequence variants of unknown clinical significance: application to BRCA1 and BRCA2. *Am J Hum Genet* **75**, 535-544, doi:10.1086/424388 (2004).
97 Li, J. *et al.* DSS1 is required for the stability of BRCA2. *Oncogene* **25**, 1186-1194, doi:10.1038/sj.onc.1209153 (2006).
98 O'connor, M. J. Targeting the DNA Damage Response in Cancer. *Mol Cell* **60**, 547-560, doi:10.1016/j.molcel.2015.10.040 (2015).

99 Livraghi, L. & Garber, J. E. PARP inhibitors in the management of breast cancer: current data and future prospects. *BMC Med* **13**, 188, doi:10.1186/s12916-015-0425-1 (2015).
100 Bryant, H. E. *et al.* Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase. *Nature* **434**, 913-917, doi:10.1038/nature03443 (2005).
101 Ranjha, L., Anand, R. & Cejka, P. The Saccharomyces cerevisiae Mlh1-Mlh3 heterodimer is an endonuclease that preferentially binds to Holliday junctions. *J Biol Chem* **289**, 5674-5686, doi:10.1074/jbc.M113.533810 (2014).
102 Simandlova, J. *et al.* FBH1 helicase disrupts RAD51 filaments in vitro and modulates homologous recombination in mammalian cells. *J Biol Chem* **288**, 34168-34180, doi:10.1074/jbc.M113.484493 (2013).
103 Munteanu, B., Braun, M. & Boonrod, K. Improvement of PCR reaction conditions for site-directed mutagenesis of big plasmids. *J Zhejiang Univ Sci B* **13**, 244-247, doi:10.1631/jzus.B1100180 (2012).
104 Cejka, P. & Kowalczykowski, S. C. The full-length Saccharomyces cerevisiae Sgs1 protein is a vigorous DNA helicase that preferentially unwinds holliday junctions. *J Biol Chem* **285**, 8290-8301, doi:10.1074/jbc.M109.083196 (2010).
105 Liu, J. & Heyer, W. D. Who's who in human recombination: BRCA2 and RAD52. *Proc Natl Acad Sci U S A* **108**, 441-442, doi:10.1073/pnas.1016614108 (2011).

5 SUPPLEMENTARY INFORMATION

5.1 List of nucleotides

Oligonucleotide	Sequence (5' to 3')
X12-3	GACGTCATAGACGATTACATTGCTAGGACATGCTGT CTAGAGACTATCGC
X12-4 C	GCGATAGTCTCTAGACAGCATGTCCTAGCAATGTAA TCGTCTATGACGTC
X12-4 SC	GCGATAGTCTCTAGACAGCATGTCCTAGCAA
X12-4 NC	GCGATAGTCTCTAGACAGCATGTCCTAGCAAGCCAGAATTCGGCAGGCTA
X12-3 TOP L	GACGTCATAGACGATTACATTGCTAGGACATGCTGTCTAGAGACTATCGCGA CTTACGTTCCATCGCTAGGTTATTTTTTTTTTTTTTTTTTT
X12-3 BOTTOM LC	AAAAAAAAAAAAAAAAAAAATAACCTAGCGATGGAACGTAAGTCGCGATAGT CTCTAGACAGCATGTCCTAGCAATGTAATCGTCTATGACGTC
X12-3 HJ1	AAAAAAAAAAAAAAAAAAAATAACCTAGCGATGGAACGTAAGTCGCGATGGG CTTAACTAGGATGCTACTGGCCCCGAATCAACCGTACTTGGG
X12-3 HJ2	CCCAAGTACGGTTGATTCGGGGCCAGTAGCATCCTAGTTAAGCCCATTACGA TTCGTTACCCATTCACTGTCAGAAGGCACCAGATAGATCTC
X12-3 HJ3	GAGATCTATCTGGTGCCTTCTGACAGTGAATGGGTAACGAATCGTAATAGTC TCTAGACAGCATGTCCTAGCAATGTAATCGTCTATGACGTC
X12-3 HJ 1S	AAAAAAAAAAAAAAAAAAAATAACCTAGCGATGGAACGTAAGTCGCGAT
X12-3 HJ2 Sa	CCCAAGTACGGTTGATTCGGGGCCAGTAGCATCCTAGTTAAGCCC
X12-3 HJ2 Sb	ATTACGATTCGTTACCCATTCACTGTCAGAAGGCACCAGATAGATCTC
PC 1253	TGGGTCAACGTGGGCAAAGATGTCCTAGCAATGTAATCGTCTATGACGTT
PC 1254	TGCCGAATTCTACCAGTGCCAGTGATGGACATCTTTGCCCACGTTGACCC
PC 1255	GTCGGATCCTCTAGACAGCTCCATGATCACTGGCACTGGTAGAATTCGGC
PC 1256	CAACGTCATAGACGATTACATTGCTACATGGAGCTGTCTAGAGGATCCGA

5.2 List of primers and sequences of MMS22L fragments

Primers design:

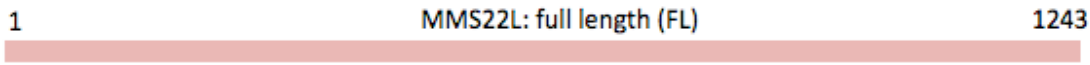
Forward: 5'-CTACTA- GGGCC''C - insert 20nt - 3' (ApaI)
Reverse: 5'-CTTCTT- C''TCGAG -TTA+ FLAG+ insert 20nt (reverse complementary) - 3' (XhoI)

FLAG tag sequence (AspTyrLysAspAspAspAspLys):
5'-gattacaaggatgacgacgataag-3'

FLAG tag (reverse complementary): 5' - CTTATCGTCGTCATCCTTGTAATC -3'

List of designed primers and MMS22L fragments sequences:

1) **Full length MMS22L (FL)**



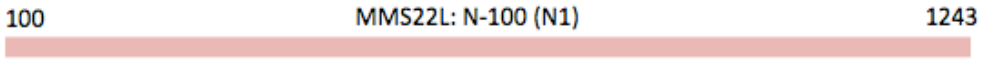
SF-P1: 5'-CTACTA GGGCCC ATGGAGAACTGTTCTGCTGC -3'

SR-P1243: 5'-CTTCTT CTCGAG TTACTTATCGTCGTCATCCTTGTAATC AGTATTATCATTTTCCAGTC -3'

ATGGAGAACTGTTCTGCTGCATCGACGTTCCTGACTGACAGCTTAGAGCTGGAGCTGGGG
ACGGAATGGTGCAAACCTCCTTACTTTTCTTGTGCTGTTGACAACAGAGGAGGAGGAAAA
CATTTTTCTGGAGAATCCTACCTCTGCAGCGGAGCCCTTAAGCGATTGATTTTGAATCTT
GACCCTTTACCAACTAATTTTGAAGAAGATACCTTGGAAATATTTGGCATTCAGTGGGTT
ACTGAAACAGCATTAGTGAATTCATCTAGAGAACTCTTTCATTTATTCAGGCAACAACTG
TACAACTTGGAAACCTTGTTACAGTCCAGTTGTGATTTTGGGAAGGTATCAACTCTACAC
TGCAAAGCAGACAATATTAGGCAGCAGTGTGTACTATTTCTCCATTATGTTAAAGTTTTC
ATCTTCAGGTATCTGAAAGTACAGAATGCTGAGAGTCATGTTCCTGTCCATCCTTATGAG
GCTTTGGAGGCTCAGCTTCCCTCAGTGTTGATTGATGAGCTTCATGGATTACTCTTGTAT
ATTGGACACCTATCTGAACTTCCCAGTGTTAATATAGGAGCATTTGTAAATCAAAACCAG
ATTAAGCTTTTTCCACCGTCATGGCATTTATTACATCTCCACTTGGATATACATTGGCTG
GTGCTAGAAATTCTTTACATGCTGGGTGAAAAATTGAAACAAGTTGTATATGGTCATCA
GTTTATGAATCTGGCAAGTGACAATTTAACCAACATCAGCCTATTTGAAGAACATTGTGA
AACTCTCCTTTGTGATTTAATAAGCCTGTCACTCAACAGGTACGACAAGGTTAGGTCTTC
TGAATCATTAATGAGTGACCAGTGTCCATGTTTATGCATTAAAGAATTATGGGTTCTACT
TATTCATCTTCTAGACCACAGAAGTAAATGGTTTGTCTCGGAATCATTTTGGAACTGGTT
GAATAAACTACTTAAAACACTGCTTGAAAAATCAAGTGACCGAAGAAGATCCTCTATGCC
TGTAATCCAGTCCAGGGATCCATTAGGTTTTAGTTGGTGGATTATTACTCATGTAGCATC
ATTTTACAAGTTTGATCGCCATGGAGTACCAGATGAAATGAGAAAAGTGGAATCAAATT
GGAACTTTGTAGAAGAACTGCTGAAAAAGTCCATCAGTGTTCAGGGTGTCATTCTAGAAG
AACAATTACGAATGTATCTTCACTGTTGTTTGACACTTTGTGATTTCTGGGAGCCAAACA

TTGCAATTGTTACCATTTTATGGGAATATTATAGTAAGAACCTGAATAGTTCCTTCAGTA
TTTCTTGGCTTCCTTTTAAAGGCCTTGCTAATACCATGAAGTCACCCTTGTCTATGCTTG
AAATGGTGAAGACTTGCTGTTGCGATAAACAAGATCAGGAACTATATAAATCCAGCAGT
AGTTATACTATTTTTCTTTGTATTCTGGCAAAAGTTGTTAAAAAAGCAATGAAGAGCAA
TGGCCCTCATCCTTGGAAACAAGTCAAAGGAAGAATATATTCAAAATTCCATCAAAAAAG
AATGGAAGAACTAACTGAAGTTGGTCTACAGAACTTTTTTAGCCTTTTTCTACTGTTAGC
AGCTGTTGCAGAGGTAGAAGATGTTGCAAGTCATGTTTTAGACCTCCTGAATTTCCTCAA
GCCTGCTTTTGTAATGTCTCAGAGAGCCCTCATTTGGAAGGGTCACATGGCCTTCCTCTT
GATGTATGCCCAGAAAAATCTGGACATTGGTGTTTTGGCTGAGAAATTTTCATGTGCTTT
CCGGGAGAAAGCAAAGGAATTCTTGGTGTCTAAGAATGAGGAAATGGTACAGAGACAGA
CTATCTGGACCCTTCTTTCCATATACATTGATGGTGTTCAAGAAGTGTTTGAGACCAGCT
ATTGCTTGTATCCTTCCCATGAAAAACTGCTTAATGATGGATTTAGTATGCTTCTGCGAG
CATGTCGAGAATCTGAACTTAGGACAGTATTGAGCTTCCTACAAGCTGTTCTGGCCAGAA
TCAGGAGTATGCATCAACAATTGTGTCAGGAACTTCAAAGGGACAATGTGGACCTATTTG
TACAGTCTTCATTATCGGCTAAAGAGCGCCACCTTGCTGCAGTTGCCAGTGCACTGTGGA
GACATTTCTTTTCATTTTTGAAGAGTCAGAGAATGTCACAGGTAGTGCCTTTCTCACAAC
TTGCGGATGCAGCTGCAGACTTTACTTTGCTAGCAATGGACATGCCAAGCACAGCTCCAT
CAGATTTTCAGCCTCAGCCAGTTATATCAATTATTCAACTTTTTGGTTGGGATGATATCA
TCTGCCCTCAAGTTGTAGCAAGATATTTAAGTCATGTCCTACAAAATAGCACATTATGTG
AAGCACTTTCTCATTCAGGCTATGTATCTTTTCAAGCCTTAACCGTAAGATCATGGATTC
GTTGTGTTTTGCAAATGTATATTAAAAACCTCTCTGGGCCTGATGATTTGCTCATAGATA
AAAATCTGGAAGAGGCAGTTGAAAAAGAGTACATGAAACAGTTGGTCAAACTGACAAGA
TTACTATTTAATCTCTCAGAAGTAAAGAGTATTTTCTCAAAGGCCCAAGCTGAATATTTA
TCCATCTCAGAAGACCCTAAAAAAGCACTTGTTCGATTCTTTGAGGCTGTTGGTGTAACT
TACGGGAACGTCCAGACACTTTCTGATAAATCTGCCATGGTCACAAAGTCCTTGGAATAC
CTTGGTGAAGTATTAAAATATATTAAGCCTTATTTGGGAAAAAAAGTTTTCAGTGCAGG
GCTGCAGCTGACTTATGGAATGATGGGAATTCTTGTGAAATCATGGGCACAAATCTTTGC
CACTTCTAAAGCCCAAAAATTACTATTCCGGATCATAGATTGTTTACTGCTGCCACATGC
AGTATTACAGCAAGAGAAGGAACTGCCTGCACCTATGTTGTCAGCAATTCAGAAAAGTCT
TCCTTTGTATCTCCAGGGCATGTGTATCGTGTGTTGTCAATCTCAAAATCCGAATGCCTA
TTTGAATCAATTGCTAGGGAATGTTATTGAGCAGTATATTGGGCGATTTCTTCCAGCTTC
ACCATATGTTTCAGATCTTGGACAACATCCTGTTTTGCTGGCATTGAGAAACACAGCCAC
TATTCCGCCAATATCATCTCTAAAGAAATGCATTGTGCAAGTCATAAGGAAATCCTACCT
TGAGTATAAGGGGTCCTCACCTCCTCCTCGCTTAGCATCCATTCTGGCCTTCATCCTCCAA
CTCTTCAAGGAAACTAACACAGACATTTATGAAGTTGAACTACTCCTCCCTGGCATTTTA
AAATGCTTGGTGTTAGTCAGTGAACCACAAGTTAAAAGGCTGGCCACAGAGAACCTGCAA
TACATGGTAAAAGCCTGCCAAGTGGGGTCAGAAGAAGAACCTTCCTCCCAGCTGACTTCT
GTGTTTAGGCAGTTTATCCAGGATTATGGTATGAGGTACTATTACCAGGTTTACAGCATT
TTAGAAACAGTAGCAACATTGGACCAGCAGGTTGTCATCCACTTGATTTCTACCCTTACT
CAGTCTCTGAAGGATTCAGAGCAGAAATGGGGCCTTGGCAGGAATATAGCACAAAGGGA
AGCCTATAGCAAACTTTTGTCTCACCTTGGACAGATGGGACAAGATGAGATGCAGAGACT
GGAAAATGATAATACT

2) **N-100 (N1)**



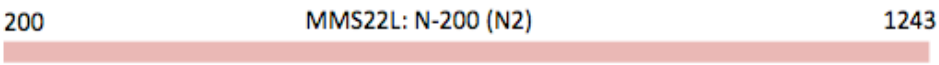
SF-N1: 5'-CTACTA GGGCCC TACAACTTGGAAACCTTGTT -3'

SR-P1243: 5'-CTTCTT CTCGAG TTACTTATCGTCGTCATCCTTGTAATC AGTATTATCATTTTCCAGTC -3'

TACAACTTGGAAACCTTGTTACAGTCCAGTTGTGATTTTGGGAAGGTATCAACTCTACAC
TGCAAAGCAGACAATATTAGGCAGCAGTGTGTACTATTTCTCCATTATGTTAAAGTTTTC
ATCTTCAGGTATCTGAAAGTACAGAATGCTGAGAGTCATGTTCCTGTCCATCCTTATGAG
GCTTTGGAGGCTCAGCTTCCCTCAGTGTTGATTGATGAGCTTCATGGATTACTCTTGTAT
ATTGGACACCTATCTGAACTTCCCAGTGTTAATATAGGAGCATTTGTAAATCAAAACCAG
ATTAAGCTTTTTCCACCGTCATGGCATTTATTACATCTCCACTTGGATATACATTGGCTG
GTGCTAGAAATTCTTTACATGCTGGGTGAAAAATTGAAACAAGTTGTATATGGTCATCA
GTTTATGAATCTGGCAAGTGACAATTTAACCAACATCAGCCTATTTGAAGAACATTGTGA
AACTCTCCTTTGTGATTTAATAAGCCTGTCACTCAACAGGTACGACAAGGTTAGGTCTTC
TGAATCATTAATGAGTGACCAGTGTCCATGTTTATGCATTAAAGAATTATGGGTTCTACT
TATTCATCTTCTAGACCACAGAAGTAAATGGTTTGTCTCGGAATCATTTTGGAACTGGTT
GAATAAACTACTTAAAACACTGCTTGAAAAATCAAGTGACCGAAGAAGATCCTCTATGCC
TGTAATCCAGTCCAGGGATCCATTAGGTTTTAGTTGGTGGATTATTACTCATGTAGCATC
ATTTTACAAGTTTGATCGCCATGGAGTACCAGATGAAATGAGAAAAGTGGAATCAAATT
GGAACTTTGTAGAAGAACTGCTGAAAAAGTCCATCAGTGTTCAGGGTGTCATTCTAGAAG
AACAATTACGAATGTATCTTCACTGTTGTTTGACACTTTGTGATTTCTGGGAGCCAAACA
TTGCAATTGTTACCATTTTATGGGAATATTATAGTAAGAACCTGAATAGTTCCTTCAGTA
TTTCTTGGCTTCCTTTTAAAGGCCTTGCTAATACCATGAAGTCACCCTTGTCTATGCTTG
AAATGGTGAAGACTTGCTGTTGCGATAAACAAGATCAGGAACTATATAAATCCAGCAGT
AGTTATACTATTTTTCTTTGTATTCTGGCAAAAGTTGTTAAAAAAGCAATGAAGAGCAA
TGGCCCTCATCCTTGGAAACAAGTCAAAGGAAGAATATATTCAAAATTCCATCAAAAAAG
AATGGAAGAACTAACTGAAGTTGGTCTACAGAACTTTTTTAGCCTTTTTCTACTGTTAGC
AGCTGTTGCAGAGGTAGAAGATGTTGCAAGTCATGTTTTAGACCTCCTGAATTTCCTCAA
GCCTGCTTTTGTAATGTCTCAGAGAGCCCTCATTTGGAAGGGTCACATGGCCTTCCTCTT
GATGTATGCCCAGAAAAATCTGGACATTGGTGTTTTGGCTGAGAAATTTTCATGTGCTTT
CCGGGAGAAAGCAAAGGAATTCTTGGTGTCTAAGAATGAGGAAATGGTACAGAGACAGA
CTATCTGGACCCTTCTTTCCATATACATTGATGGTGTTCAAGAAGTGTTTGAGACCAGCT
ATTGCTTGTATCCTTCCCATGAAAAACTGCTTAATGATGGATTTAGTATGCTTCTGCGAG
CATGTCGAGAATCTGAACTTAGGACAGTATTGAGCTTCCTACAAGCTGTTCTGGCCAGAA
TCAGGAGTATGCATCAACAATTGTGTCAGGAACTTCAAAGGGACAATGTGGACCTATTTG
TACAGTCTTCATTATCGGCTAAAGAGCGCCACCTTGCTGCAGTTGCCAGTGCACTGTGGA
GACATTTCTTTTCATTTTTGAAGAGTCAGAGAATGTCACAGGTAGTGCCTTTCTCACAAC
TTGCGGATGCAGCTGCAGACTTTACTTTGCTAGCAATGGACATGCCAAGCACAGCTCCAT
CAGATTTTCAGCCTCAGCCAGTTATATCAATTATTCAACTTTTTGGTTGGGATGATATCA
TCTGCCCTCAAGTTGTAGCAAGATATTTAAGTCATGTCCTACAAAATAGCACATTATGTG
AAGCACTTTCTCATTCAGGCTATGTATCTTTTCAAGCCTTAACCGTAAGATCATGGATTC
GTTGTGTTTTGCAAATGTATATTAAAAACCTCTCTGGGCCTGATGATTTGCTCATAGATA
AAAATCTGGAAGAGGCAGTTGAAAAAGAGTACATGAAACAGTTGGTCAAACTGACAAGA
TTACTATTTAATCTCTCAGAAGTAAAGAGTATTTTCTCAAAGGCCCAAGCTGAATATTTA

TCCATCTCAGAAGACCCTAAAAAAGCACTTGTTCGATTCTTTGAGGCTGTTGGTGTAACT
TACGGGAACGTCCAGACACTTTCTGATAAATCTGCCATGGTCACAAAGTCCTTGGAATAC
CTTGGTGAAGTATTAAAATATATTAAGCCTTATTTGGGAAAAAAAGTTTTCAGTGCAGG
GCTGCAGCTGACTTATGGAATGATGGGAATTCTTGTGAAATCATGGGCACAAATCTTTGC
CACTTCTAAAGCCCAAAAATTACTATTCCGGATCATAGATTGTTTACTGCTGCCACATGC
AGTATTACAGCAAGAGAAGGAACTGCCTGCACCTATGTTGTCAGCAATTCAGAAAAGTCT
TCCTTTGTATCTCCAGGGCATGTGTATCGTGTGTTGTCAATCTCAAAATCCGAATGCCTA
TTTGAATCAATTGCTAGGGAATGTTATTGAGCAGTATATTGGGCGATTTCTTCCAGCTTC
ACCATATGTTTCAGATCTTGGACAACATCCTGTTTTGCTGGCATTGAGAAACACAGCCAC
TATTCCGCCAATATCATCTCTAAAGAAATGCATTGTGCAAGTCATAAGGAAATCCTACCT
TGAGTATAAGGGGTCCTCACCTCCTCCTCGCTTAGCATCCATTCTGGCCTTCATCCTCCAA
CTCTTCAAGGAAACTAACACAGACATTTATGAAGTTGAACTACTCCTCCCTGGCATTTTA
AAATGCTTGGTGTTAGTCAGTGAACCACAAGTTAAAAGGCTGGCCACAGAGAACCTGCAA
TACATGGTAAAAGCCTGCCAAGTGGGGTCAGAAGAAGAACCTTCCTCCCAGCTGACTTCT
GTGTTTAGGCAGTTTATCCAGGATTATGGTATGAGGTACTATTACCAGGTTTACAGCATT
TTAGAAACAGTAGCAACATTGGACCAGCAGGTTGTCATCCACTTGATTTCTACCCTTACT
CAGTCTCTGAAGGATTCAGAGCAGAAATGGGGCCTTGGCAGGAATATAGCACAAAGGGA
AGCCTATAGCAAACTTTTGTCTCACCTTGGACAGATGGGACAAGATGAGATGCAGAGACT
GGAAAATGATAATACT

3) **N-200 (N2)**



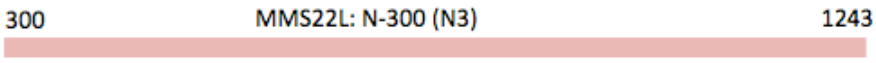
SF-N2: 5'-CTACTA GGGCCC ATTAAGCTTTTTCCACCGTC-3'

SR-P1243: 5'-CTTCTT CTCGAG TTACTTATCGTCGTCATCCTTGTAATC
AGTATTATCATTTTCCAGTC -3'

ATTAAGCTTTTTCCACCGTCATGGCATTTATTACATCTCCACTTGGATATACATTGGCTG
GTGCTAGAAATTCTTTACATGCTGGGTGAAAAATTGAAACAAGTTGTATATGGTCATCA
GTTTATGAATCTGGCAAGTGACAATTTAACCAACATCAGCCTATTTGAAGAACATTGTGA
AACTCTCCTTTGTGATTTAATAAGCCTGTCACTCAACAGGTACGACAAGGTTAGGTCTTC
TGAATCATTAATGAGTGACCAGTGTCCATGTTTATGCATTAAAGAATTATGGGTTCTACT
TATTCATCTTCTAGACCACAGAAGTAAATGGTTTGTCTCGGAATCATTTTGGAACTGGTT
GAATAAACTACTTAAAACACTGCTTGAAAAATCAAGTGACCGAAGAAGATCCTCTATGCC
TGTAATCCAGTCCAGGGATCCATTAGGTTTTAGTTGGTGGATTATTACTCATGTAGCATC
ATTTTACAAGTTTGATCGCCATGGAGTACCAGATGAAATGAGAAAAGTGGAATCAAATT
GGAACTTTGTAGAAGAACTGCTGAAAAAGTCCATCAGTGTTCAGGGTGTCATTCTAGAAG
AACAATTACGAATGTATCTTCACTGTTGTTTGACACTTTGTGATTTCTGGGAGCCAAACA
TTGCAATTGTTACCATTTTATGGGAATATTATAGTAAGAACCTGAATAGTTCCTTCAGTA
TTTCTTGGCTTCCTTTTAAAGGCCTTGCTAATACCATGAAGTCACCCTTGTCTATGCTTG
AAATGGTGAAGACTTGCTGTTGCGATAAACAAGATCAGGAACTATATAAATCCAGCAGT
AGTTATACTATTTTTCTTTGTATTCTGGCAAAAGTTGTTAAAAAAGCAATGAAGAGCAA
TGGCCCTCATCCTTGGAAACAAGTCAAAGGAAGAATATATTCAAAATTCCATCAAAAAAG
AATGGAAGAACTAACTGAAGTTGGTCTACAGAACTTTTTTAGCCTTTTTCTACTGTTAGC

AGCTGTTGCAGAGGTAGAAGATGTTGCAAGTCATGTTTTAGACCTCCTGAATTTCCTCAA
GCCTGCTTTTGTAATGTCTCAGAGAGCCCTCATTTGGAAGGGTCACATGGCCTTCCTCTT
GATGTATGCCCAGAAAAATCTGGACATTGGTGTTTTGGCTGAGAAATTTTCATGTGCTTT
CCGGGAGAAAGCAAAGGAATTCTTGGTGTCTAAGAATGAGGAAATGGTACAGAGACAGA
CTATCTGGACCCTTCTTTCCATATACATTGATGGTGTTCAAGAAGTGTTTGAGACCAGCT
ATTGCTTGTATCCTTCCCATGAAAAACTGCTTAATGATGGATTTAGTATGCTTCTGCGAG
CATGTCGAGAATCTGAACTTAGGACAGTATTGAGCTTCCTACAAGCTGTTCTGGCCAGAA
TCAGGAGTATGCATCAACAATTGTGTCAGGAACTTCAAAGGGACAATGTGGACCTATTTG
TACAGTCTTCATTATCGGCTAAAGAGCGCCACCTTGCTGCAGTTGCCAGTGCACTGTGGA
GACATTTCTTTTCATTTTTGAAGAGTCAGAGAATGTCACAGGTAGTGCCTTTCTCACAAC
TTGCGGATGCAGCTGCAGACTTTACTTTGCTAGCAATGGACATGCCAAGCACAGCTCCAT
CAGATTTTCAGCCTCAGCCAGTTATATCAATTATTCAACTTTTTGGTTGGGATGATATCA
TCTGCCCTCAAGTTGTAGCAAGATATTTAAGTCATGTCCTACAAAATAGCACATTATGTG
AAGCACTTTCTCATTCAGGCTATGTATCTTTTCAAGCCTTAACCGTAAGATCATGGATTC
GTTGTGTTTTGCAAATGTATATTAAAAACCTCTCTGGGCCTGATGATTTGCTCATAGATA
AAAATCTGGAAGAGGCAGTTGAAAAAGAGTACATGAAACAGTTGGTCAAACTGACAAGA
TTACTATTTAATCTCTCAGAAGTAAAGAGTATTTTCTCAAAGGCCCAAGCTGAATATTTA
TCCATCTCAGAAGACCCTAAAAAAGCACTTGTTCGATTCTTTGAGGCTGTTGGTGTAACT
TACGGGAACGTCCAGACACTTTCTGATAAATCTGCCATGGTCACAAAGTCCTTGGAATAC
CTTGGTGAAGTATTAAAATATATTAAGCCTTATTTGGGAAAAAAAGTTTTCAGTGCAGG
GCTGCAGCTGACTTATGGAATGATGGGAATTCTTGTGAAATCATGGGCACAAATCTTTGC
CACTTCTAAAGCCCAAAAATTACTATTCCGGATCATAGATTGTTTACTGCTGCCACATGC
AGTATTACAGCAAGAGAAGGAACTGCCTGCACCTATGTTGTCAGCAATTCAGAAAAGTCT
TCCTTTGTATCTCCAGGGCATGTGTATCGTGTGTTGTCAATCTCAAAATCCGAATGCCTA
TTTGAATCAATTGCTAGGGAATGTTATTGAGCAGTATATTGGGCGATTTCTTCCAGCTTC
ACCATATGTTTCAGATCTTGGACAACATCCTGTTTTGCTGGCATTGAGAAACACAGCCAC
TATTCCGCCAATATCATCTCTAAAGAAATGCATTGTGCAAGTCATAAGGAAATCCTACCT
TGAGTATAAGGGGTCCTCACCTCCTCCTCGCTTAGCATCCATTCTGGCCTTCATCCTCCAA
CTCTTCAAGGAAACTAACACAGACATTTATGAAGTTGAACTACTCCTCCCTGGCATTTTA
AAATGCTTGGTGTTAGTCAGTGAACCACAAGTTAAAAGGCTGGCCACAGAGAACCTGCAA
TACATGGTAAAAGCCTGCCAAGTGGGGTCAGAAGAAGAACCTTCCTCCCAGCTGACTTCT
GTGTTTAGGCAGTTTATCCAGGATTATGGTATGAGGTACTATTACCAGGTTTACAGCATT
TTAGAAACAGTAGCAACATTGGACCAGCAGGTTGTCATCCACTTGATTTCTACCCTTACT
CAGTCTCTGAAGGATTCAGAGCAGAAATGGGGCCTTGGCAGGAATATAGCACAAAGGGA
AGCCTATAGCAAACTTTTGTCTCACCTTGGACAGATGGGACAAGATGAGATGCAGAGACT
GGAAAATGATAATACT

4) **N-300 (N3)**

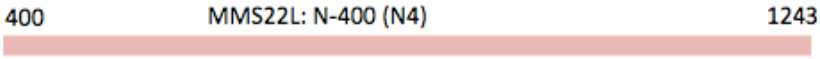


SF-N3: 5'-CTACTA GGGCCC ATTCATCTTCTAGACCACAG-3'

SR-P1243: 5'-CTTCTT CTCGAG TTACTTATCGTCGTCATCCTTGTAATC
AGTATTATCATTTTCCAGTC -3'

ATTCATCTTCTAGACCACAGAAGTAAATGGTTTGTCTCGGAATCATTTTGGAACTGGTTG
AATAAACTACTTAAAACACTGCTTGAAAAATCAAGTGACCGAAGAAGATCCTCTATGCCT
GTAATCCAGTCCAGGGATCCATTAGGTTTTAGTTGGTGGATTATTACTCATGTAGCATCA
TTTTACAAGTTTGATCGCCATGGAGTACCAGATGAAATGAGAAAAGTGGAATCAAATTG
GAACTTTGTAGAAGAACTGCTGAAAAAGTCCATCAGTGTTCAGGGTGTCATTCTAGAAGA
ACAATTACGAATGTATCTTCACTGTTGTTTGACACTTTGTGATTTCTGGGAGCCAAACAT
TGCAATTGTTACCATTTTATGGGAATATTATAGTAAGAACCTGAATAGTTCCTTCAGTAT
TTCTTGGCTTCCTTTTAAAGGCCTTGCTAATACCATGAAGTCACCCTTGTCTATGCTTGA
AATGGTGAAGACTTGCTGTTGCGATAAACAAGATCAGGAACTATATAAATCCAGCAGTA
GTTATACTATTTTTCTTTGTATTCTGGCAAAAGTTGTTAAAAAAGCAATGAAGAGCAAT
GGCCCTCATCCTTGGAAACAAGTCAAAGGAAGAATATATTCAAAATTCCATCAAAAAAGA
ATGGAAGAACTAACTGAAGTTGGTCTACAGAACTTTTTTAGCCTTTTTCTACTGTTAGCA
GCTGTTGCAGAGGTAGAAGATGTTGCAAGTCATGTTTTAGACCTCCTGAATTTCCTCAAG
CCTGCTTTTGTAATGTCTCAGAGAGCCCTCATTTGGAAGGGTCACATGGCCTTCCTCTTG
ATGTATGCCCAGAAAAATCTGGACATTGGTGTTTTGGCTGAGAAATTTTCATGTGCTTTC
CGGGAGAAAGCAAAGGAATTCTTGGTGTCTAAGAATGAGGAAATGGTACAGAGACAGAC
TATCTGGACCCTTCTTTCCATATACATTGATGGTGTTCAAGAAGTGTTTGAGACCAGCTA
TTGCTTGTATCCTTCCCATGAAAAACTGCTTAATGATGGATTTAGTATGCTTCTGCGAGC
ATGTCGAGAATCTGAACTTAGGACAGTATTGAGCTTCCTACAAGCTGTTCTGGCCAGAAT
CAGGAGTATGCATCAACAATTGTGTCAGGAACTTCAAAGGGACAATGTGGACCTATTTGT
ACAGTCTTCATTATCGGCTAAAGAGCGCCACCTTGCTGCAGTTGCCAGTGCACTGTGGAG
ACATTTCTTTTCATTTTTGAAGAGTCAGAGAATGTCACAGGTAGTGCCTTTCTCACAACT
TGCGGATGCAGCTGCAGACTTTACTTTGCTAGCAATGGACATGCCAAGCACAGCTCCATC
AGATTTTCAGCCTCAGCCAGTTATATCAATTATTCAACTTTTTGGTTGGGATGATATCAT
CTGCCCTCAAGTTGTAGCAAGATATTTAAGTCATGTCCTACAAAATAGCACATTATGTGA
AGCACTTTCTCATTCAGGCTATGTATCTTTTCAAGCCTTAACCGTAAGATCATGGATTCG
TTGTGTTTTGCAAATGTATATTAAAAACCTCTCTGGGCCTGATGATTTGCTCATAGATAA
AAATCTGGAAGAGGCAGTTGAAAAAGAGTACATGAAACAGTTGGTCAAACTGACAAGAT
TACTATTTAATCTCTCAGAAGTAAAGAGTATTTTCTCAAAGGCCCAAGCTGAATATTTAT
CCATCTCAGAAGACCCTAAAAAAGCACTTGTTCGATTCTTTGAGGCTGTTGGTGTAACTT
ACGGGAACGTCCAGACACTTTCTGATAAATCTGCCATGGTCACAAAGTCCTTGGAATACC
TTGGTGAAGTATTAAAATATATTAAGCCTTATTTGGGAAAAAAAGTTTTCAGTGCAGGG
CTGCAGCTGACTTATGGAATGATGGGAATTCTTGTGAAATCATGGGCACAAATCTTTGCC
ACTTCTAAAGCCCAAAAATTACTATTCCGGATCATAGATTGTTTACTGCTGCCACATGCA
GTATTACAGCAAGAGAAGGAACTGCCTGCACCTATGTTGTCAGCAATTCAGAAAAGTCTT
CCTTTGTATCTCCAGGGCATGTGTATCGTGTGTTGTCAATCTCAAAATCCGAATGCCTAT
TTGAATCAATTGCTAGGGAATGTTATTGAGCAGTATATTGGGCGATTTCTTCCAGCTTCA
CCATATGTTTCAGATCTTGGACAACATCCTGTTTTGCTGGCATTGAGAAACACAGCCACT
ATTCCGCCAATATCATCTCTAAAGAAATGCATTGTGCAAGTCATAAGGAAATCCTACCTT
GAGTATAAGGGGTCCTCACCTCCTCCTCGCTTAGCATCCATTCTGGCCTTCATCCTCCAAC
TCTTCAAGGAAACTAACACAGACATTTATGAAGTTGAACTACTCCTCCCTGGCATTTTAA
AATGCTTGGTGTTAGTCAGTGAACCACAAGTTAAAAGGCTGGCCACAGAGAACCTGCAAT
ACATGGTAAAAGCCTGCCAAGTGGGGTCAGAAGAAGAACCTTCCTCCCAGCTGACTTCTG
TGTTTAGGCAGTTTATCCAGGATTATGGTATGAGGTACTATTACCAGGTTTACAGCATTT
TAGAAACAGTAGCAACATTGGACCAGCAGGTTGTCATCCACTTGATTTCTACCCTTACTC
AGTCTCTGAAGGATTCAGAGCAGAAATGGGGCCTTGGCAGGAATATAGCACAAAGGGAA
GCCTATAGCAAACTTTTGTCTCACCTTGGACAGATGGGACAAGATGAGATGCAGAGACTG
GAAAATGATAATACT

5) **N-400 (N4)**



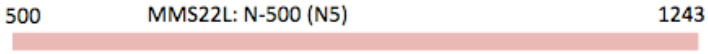
SF-N4: 5'-CTACTA GGGCCC CAATTACGAATGTATCTTCA-3'

SR-P1243: 5'-CTTCTT CTCGAG TTACTTATCGTCGTCATCCTTGTAATC
AGTATTATCATTTTCCAGTC -3'

CAATTACGAATGTATCTTCACTGTTGTTTGACACTTTGTGATTTCTGGGAGCCAAACATT
GCAATTGTTACCATTTTATGGGAATATTATAGTAAGAACCTGAATAGTTCCTTCAGTATT
TCTTGGCTTCCTTTTAAAGGCCTTGCTAATACCATGAAGTCACCCTTGTCTATGCTTGAA
ATGGTGAAGACTTGCTGTTGCGATAAACAAGATCAGGAACTATATAAATCCAGCAGTAG
TTATACTATTTTTCTTTGTATTCTGGCAAAAGTTGTTAAAAAAGCAATGAAGAGCAATG
GCCCTCATCCTTGGAAACAAGTCAAAGGAAGAATATATTCAAAATTCCATCAAAAAAGAA
TGGAAGAACTAACTGAAGTTGGTCTACAGAACTTTTTTAGCCTTTTTCTACTGTTAGCAG
CTGTTGCAGAGGTAGAAGATGTTGCAAGTCATGTTTTAGACCTCCTGAATTTCCTCAAGC
CTGCTTTTGTAATGTCTCAGAGAGCCCTCATTTGGAAGGGTCACATGGCCTTCCTCTTGA
TGTATGCCCAGAAAAATCTGGACATTGGTGTTTTGGCTGAGAAATTTTCATGTGCTTTCC
GGGAGAAAGCAAAGGAATTCTTGGTGTCTAAGAATGAGGAAATGGTACAGAGACAGACT
ATCTGGACCCTTCTTTCCATATACATTGATGGTGTTCAAGAAGTGTTTGAGACCAGCTAT
TGCTTGTATCCTTCCCATGAAAAACTGCTTAATGATGGATTTAGTATGCTTCTGCGAGCA
TGTCGAGAATCTGAACTTAGGACAGTATTGAGCTTCCTACAAGCTGTTCTGGCCAGAATC
AGGAGTATGCATCAACAATTGTGTCAGGAACTTCAAAGGGACAATGTGGACCTATTTGTA
CAGTCTTCATTATCGGCTAAAGAGCGCCACCTTGCTGCAGTTGCCAGTGCACTGTGGAGA
CATTTCTTTTCATTTTTGAAGAGTCAGAGAATGTCACAGGTAGTGCCTTTCTCACAACTT
GCGGATGCAGCTGCAGACTTTACTTTGCTAGCAATGGACATGCCAAGCACAGCTCCATCA
GATTTTCAGCCTCAGCCAGTTATATCAATTATTCAACTTTTTGGTTGGGATGATATCATC
TGCCCTCAAGTTGTAGCAAGATATTTAAGTCATGTCCTACAAAATAGCACATTATGTGAA
GCACTTTCTCATTCAGGCTATGTATCTTTTCAAGCCTTAACCGTAAGATCATGGATTCGT
TGTGTTTTGCAAATGTATATTAAAAACCTCTCTGGGCCTGATGATTTGCTCATAGATAAA
AATCTGGAAGAGGCAGTTGAAAAAGAGTACATGAAACAGTTGGTCAAACTGACAAGATT
ACTATTTAATCTCTCAGAAGTAAAGAGTATTTTCTCAAAGGCCCAAGCTGAATATTTATC
CATCTCAGAAGACCCTAAAAAAGCACTTGTTCGATTCTTTGAGGCTGTTGGTGTAACTTA
CGGGAACGTCCAGACACTTTCTGATAAATCTGCCATGGTCACAAAGTCCTTGGAATACCT
TGGTGAAGTATTAAAATATATTAAGCCTTATTTGGGAAAAAAAGTTTTCAGTGCAGGGC
TGCAGCTGACTTATGGAATGATGGGAATTCTTGTGAAATCATGGGCACAAATCTTTGCCA
CTTCTAAAGCCCAAAAATTACTATTCCGGATCATAGATTGTTTACTGCTGCCACATGCAG
TATTACAGCAAGAGAAGGAACTGCCTGCACCTATGTTGTCAGCAATTCAGAAAAGTCTTC
CTTTGTATCTCCAGGGCATGTGTATCGTGTGTTGTCAATCTCAAAATCCGAATGCCTATT
TGAATCAATTGCTAGGGAATGTTATTGAGCAGTATATTGGGCGATTTCTTCCAGCTTCAC
CATATGTTTCAGATCTTGGACAACATCCTGTTTTGCTGGCATTGAGAAACACAGCCACTA
TTCCGCCAATATCATCTCTAAAGAAATGCATTGTGCAAGTCATAAGGAAATCCTACCTTG
AGTATAAGGGGTCCTCACCTCCTCCTCGCTTAGCATCCATTCTGGCCTTCATCCTCCAACT
CTTCAAGGAAACTAACACAGACATTTATGAAGTTGAACTACTCCTCCCTGGCATTTTAAA
ATGCTTGGTGTTAGTCAGTGAACCACAAGTTAAAAGGCTGGCCACAGAGAACCTGCAATA
CATGGTAAAAGCCTGCCAAGTGGGGTCAGAAGAAGAACCTTCCTCCCAGCTGACTTCTGT

GTTTAGGCAGTTTATCCAGGATTATGGTATGAGGTACTATTACCAGGTTTACAGCATTTT
AGAAACAGTAGCAACATTGGACCAGCAGGTTGTCATCCACTTGATTTCTACCCTTACTCA
GTCTCTGAAGGATTCAGAGCAGAAATGGGGCCTTGGCAGGAATATAGCACAAAGGGAAG
CCTATAGCAAACTTTTGTCTCACCTTGGACAGATGGGACAAGATGAGATGCAGAGACTGG
AAAATGATAATACT

6) **N-500 (N5)**



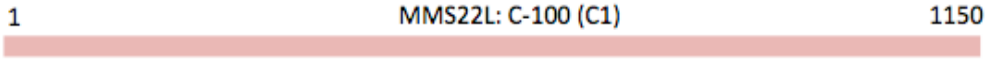
SF-N5: 5'-CTACTA GGGCCCCCTCATCCTTGGAAACAAGT-3'

SR-P1243: 5'-CTTCTT CTCGAG TTACTTATCGTCGTCATCCTTGTAATC
AGTATTATCATTTTCCAGTC -3'

CCTCATCCTTGGAAACAAGTCAAAGGAAGAATATATTCAAAATTCCATCAAAAAAGAAT
GGAAGAACTAACTGAAGTTGGTCTACAGAACTTTTTTAGCCTTTTTCTACTGTTAGCAGC
TGTTGCAGAGGTAGAAGATGTTGCAAGTCATGTTTTAGACCTCCTGAATTTCCTCAAGCC
TGCTTTTGTAATGTCTCAGAGAGCCCTCATTTGGAAGGGTCACATGGCCTTCCTCTTGAT
GTATGCCCAGAAAAATCTGGACATTGGTGTTTTGGCTGAGAAATTTTCATGTGCTTTCCG
GGAGAAAGCAAAGGAATTCTTGGTGTCTAAGAATGAGGAAATGGTACAGAGACAGACTA
TCTGGACCCTTCTTTCCATATACATTGATGGTGTTCAAGAAGTGTTTGAGACCAGCTATT
GCTTGTATCCTTCCCATGAAAAACTGCTTAATGATGGATTTAGTATGCTTCTGCGAGCAT
GTCGAGAATCTGAACTTAGGACAGTATTGAGCTTCCTACAAGCTGTTCTGGCCAGAATCA
GGAGTATGCATCAACAATTGTGTCAGGAACTTCAAAGGGACAATGTGGACCTATTTGTAC
AGTCTTCATTATCGGCTAAAGAGCGCCACCTTGCTGCAGTTGCCAGTGCACTGTGGAGAC
ATTTCTTTTCATTTTTGAAGAGTCAGAGAATGTCACAGGTAGTGCCTTTCTCACAACTTG
CGGATGCAGCTGCAGACTTTACTTTGCTAGCAATGGACATGCCAAGCACAGCTCCATCAG
ATTTTCAGCCTCAGCCAGTTATATCAATTATTCAACTTTTTGGTTGGGATGATATCATCT
GCCCTCAAGTTGTAGCAAGATATTTAAGTCATGTCCTACAAAATAGCACATTATGTGAAG
CACTTTCTCATTCAGGCTATGTATCTTTTCAAGCCTTAACCGTAAGATCATGGATTCGTT
GTGTTTTGCAAATGTATATTAAAAACCTCTCTGGGCCTGATGATTTGCTCATAGATAAAA
ATCTGGAAGAGGCAGTTGAAAAAGAGTACATGAAACAGTTGGTCAAACTGACAAGATTA
CTATTTAATCTCTCAGAAGTAAAGAGTATTTTCTCAAAGGCCCAAGCTGAATATTTATCC
ATCTCAGAAGACCCTAAAAAAGCACTTGTTCGATTCTTTGAGGCTGTTGGTGTAACTTAC
GGGAACGTCCAGACACTTTCTGATAAATCTGCCATGGTCACAAAGTCCTTGGAATACCTT
GGTGAAGTATTAAAATATATTAAGCCTTATTTGGGAAAAAAAGTTTTCAGTGCAGGGCT
GCAGCTGACTTATGGAATGATGGGAATTCTTGTGAAATCATGGGCACAAATCTTTGCCAC
TTCTAAAGCCCAAAAATTACTATTCCGGATCATAGATTGTTTACTGCTGCCACATGCAGT
ATTACAGCAAGAGAAGGAACTGCCTGCACCTATGTTGTCAGCAATTCAGAAAAGTCTTCC
TTTGTATCTCCAGGGCATGTGTATCGTGTGTTGTCAATCTCAAAATCCGAATGCCTATTT
GAATCAATTGCTAGGGAATGTTATTGAGCAGTATATTGGGCGATTTCTTCCAGCTTCACC
ATATGTTTCAGATCTTGGACAACATCCTGTTTTGCTGGCATTGAGAAACACAGCCACTAT
TCCGCCAATATCATCTCTAAAGAAATGCATTGTGCAAGTCATAAGGAAATCCTACCTTGA
GTATAAGGGGTCCTCACCTCCTCCTCGCTTAGCATCCATTCTGGCCTTCATCCTCCAACTC

TTCAAGGAAACTAACACAGACATTTATGAAGTTGAACTACTCCTCCCTGGCATTTTAAAA
TGCTTGGTGTTAGTCAGTGAACCACAAGTTAAAAGGCTGGCCACAGAGAACCTGCAATAC
ATGGTAAAAGCCTGCCAAGTGGGGTCAGAAGAAGAACCTTCCTCCCAGCTGACTTCTGTG
TTTAGGCAGTTTATCCAGGATTATGGTATGAGGTACTATTACCAGGTTTACAGCATTTTA
GAAACAGTAGCAACATTGGACCAGCAGGTTGTCATCCACTTGATTTCTACCCTTACTCAG
TCTCTGAAGGATTCAGAGCAGAAATGGGGCCTTGGCAGGAATATAGCACAAAGGGAAGCC
TATAGCAAACTTTTGTCTCACCTTGGACAGATGGGACAAGATGAGATGCAGAGACTGGAA
AATGATAATACT

7) **C-100 (C1)**



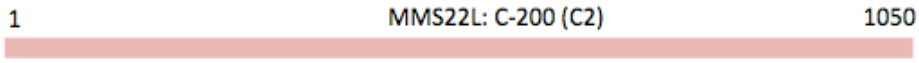
SF-P1: 5'-CTACTA GGGCCC ATGGAGAACTGTTCTGCTGC -3'

SR-C1: 5'-CTTCTT CTCGAG TTACTTATCGTCGTCATCCTTGTAATC
TTCTTCTGACCCCACTTGGC -3'

ATGGAGAACTGTTCTGCTGCATCGACGTTCCTGACTGACAGCTTAGAGCTGGAGCTGGGG
ACGGAATGGTGCAAACCTCCTTACTTTTCTTGTGCTGTTGACAACAGAGGAGGAGGAAAA
CATTTTTCTGGAGAATCCTACCTCTGCAGCGGAGCCCTTAAGCGATTGATTTTGAATCTT
GACCCTTTACCAACTAATTTTGAAGAAGATACCTTGGAAATATTTGGCATTCAGTGGGTT
ACTGAAACAGCATTAGTGAATTCATCTAGAGAACTCTTTCATTTATTCAGGCAACAACTG
TACAACTTGGAAACCTTGTTACAGTCCAGTTGTGATTTTGGGAAGGTATCAACTCTACAC
TGCAAAGCAGACAATATTAGGCAGCAGTGTGTACTATTTCTCCATTATGTTAAAGTTTTC
ATCTTCAGGTATCTGAAAGTACAGAATGCTGAGAGTCATGTTCCTGTCCATCCTTATGAG
GCTTTGGAGGCTCAGCTTCCCTCAGTGTTGATTGATGAGCTTCATGGATTACTCTTGTAT
ATTGGACACCTATCTGAACTTCCCAGTGTTAATATAGGAGCATTTGTAAATCAAAACCAG
ATTAAGCTTTTTCCACCGTCATGGCATTTATTACATCTCCACTTGGATATACATTGGCTG
GTGCTAGAAATTCTTTACATGCTGGGTGAAAAATTGAAACAAGTTGTATATGGTCATCA
GTTTATGAATCTGGCAAGTGACAATTTAACCAACATCAGCCTATTTGAAGAACATTGTGA
AACTCTCCTTTGTGATTTAATAAGCCTGTCACTCAACAGGTACGACAAGGTTAGGTCTTC
TGAATCATTAATGAGTGACCAGTGTCCATGTTTATGCATTAAAGAATTATGGGTTCTACT
TATTCATCTTCTAGACCACAGAAGTAAATGGTTTGTCTCGGAATCATTTTGGAACTGGTT
GAATAAACTACTTAAAACACTGCTTGAAAAATCAAGTGACCGAAGAAGATCCTCTATGCC
TGTAATCCAGTCCAGGGATCCATTAGGTTTTAGTTGGTGGATTATTACTCATGTAGCATC
ATTTTACAAGTTTGATCGCCATGGAGTACCAGATGAAATGAGAAAAGTGGAATCAAATT
GGAACTTTGTAGAAGAACTGCTGAAAAAGTCCATCAGTGTTCAGGGTGTCATTCTAGAAG
AACAATTACGAATGTATCTTCACTGTTGTTTGACACTTTGTGATTTCTGGGAGCCAAACA
TTGCAATTGTTACCATTTTATGGGAATATTATAGTAAGAACCTGAATAGTTCCTTCAGTA
TTTCTTGGCTTCCTTTTAAAGGCCTTGCTAATACCATGAAGTCACCCTTGTCTATGCTTG
AAATGGTGAAGACTTGCTGTTGCGATAAACAAGATCAGGAACTATATAAATCCAGCAGT
AGTTATACTATTTTTCTTTGTATTCTGGCAAAAGTTGTTAAAAAAGCAATGAAGAGCAA
TGGCCCTCATCCTTGGAAACAAGTCAAAGGAAGAATATATTCAAAATTCCATCAAAAAAG
AATGGAAGAACTAACTGAAGTTGGTCTACAGAACTTTTTTAGCCTTTTTCTACTGTTAGC
AGCTGTTGCAGAGGTAGAAGATGTTGCAAGTCATGTTTTAGACCTCCTGAATTTCCTCAA

GCCTGCTTTTGTAATGTCTCAGAGAGCCCTCATTTGGAAGGGTCACATGGCCTTCCTCTT
GATGTATGCCCAGAAAAATCTGGACATTGGTGTTTTGGCTGAGAAATTTTCATGTGCTTT
CCGGGAGAAAGCAAAGGAATTCTTGGTGTCTAAGAATGAGGAAATGGTACAGAGACAGA
CTATCTGGACCCTTCTTTCCATATACATTGATGGTGTTCAAGAAGTGTTTGAGACCAGCT
ATTGCTTGTATCCTTCCCATGAAAAACTGCTTAATGATGGATTTAGTATGCTTCTGCGAG
CATGTCGAGAATCTGAACTTAGGACAGTATTGAGCTTCCTACAAGCTGTTCTGGCCAGAA
TCAGGAGTATGCATCAACAATTGTGTCAGGAACTTCAAAGGGACAATGTGGACCTATTTG
TACAGTCTTCATTATCGGCTAAAGAGCGCCACCTTGCTGCAGTTGCCAGTGCACTGTGGA
GACATTTCTTTTCATTTTTGAAGAGTCAGAGAATGTCACAGGTAGTGCCTTTCTCACAAC
TTGCGGATGCAGCTGCAGACTTTACTTTGCTAGCAATGGACATGCCAAGCACAGCTCCAT
CAGATTTTCAGCCTCAGCCAGTTATATCAATTATTCAACTTTTTGGTTGGGATGATATCA
TCTGCCCTCAAGTTGTAGCAAGATATTTAAGTCATGTCCTACAAAATAGCACATTATGTG
AAGCACTTTCTCATTCAGGCTATGTATCTTTTCAAGCCTTAACCGTAAGATCATGGATTC
GTTGTGTTTTGCAAATGTATATTAAAAACCTCTCTGGGCCTGATGATTTGCTCATAGATA
AAAATCTGGAAGAGGCAGTTGAAAAAGAGTACATGAAACAGTTGGTCAAACTGACAAGA
TTACTATTTAATCTCTCAGAAGTAAAGAGTATTTTCTCAAAGGCCCAAGCTGAATATTTA
TCCATCTCAGAAGACCCTAAAAAAGCACTTGTTCGATTCTTTGAGGCTGTTGGTGTAACT
TACGGGAACGTCCAGACACTTTCTGATAAATCTGCCATGGTCACAAAGTCCTTGGAATAC
CTTGGTGAAGTATTAAAATATATTAAGCCTTATTTGGGAAAAAAAGTTTTCAGTGCAGG
GCTGCAGCTGACTTATGGAATGATGGGAATTCTTGTGAAATCATGGGCACAAATCTTTGC
CACTTCTAAAGCCCAAAAATTACTATTCCGGATCATAGATTGTTTACTGCTGCCACATGC
AGTATTACAGCAAGAGAAGGAACTGCCTGCACCTATGTTGTCAGCAATTCAGAAAAGTCT
TCCTTTGTATCTCCAGGGCATGTGTATCGTGTGTTGTCAATCTCAAAATCCGAATGCCTA
TTTGAATCAATTGCTAGGGAATGTTATTGAGCAGTATATTGGGCGATTTCTTCCAGCTTC
ACCATATGTTTCAGATCTTGGACAACATCCTGTTTTGCTGGCATTGAGAAACACAGCCAC
TATTCCGCCAATATCATCTCTAAAGAAATGCATTGTGCAAGTCATAAGGAAATCCTACCT
TGAGTATAAGGGGTCCTCACCTCCTCCTCGCTTAGCATCCATTCTGGCCTTCATCCTCCAA
CTCTTCAAGGAAACTAACACAGACATTTATGAAGTTGAACTACTCCTCCCTGGCATTTTA
AAATGCTTGGTGTTAGTCAGTGAACCACAAGTTAAAAGGCTGGCCACAGAGAACCTGCAA
TACATGGTAAAAGCCTGCCAAGTGGGGTCAGAAGAA

8) **C-200 (C2)**



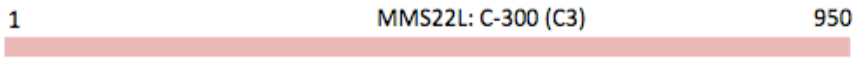
SF-P1: 5'-CTACTA GGGCCC ATGGAGAACTGTTCTGCTGC -3'

SR-P1050: 5'-CTTCTT CTCGAG TTACTTATCGTCGTCATCCTTGTAATC
CAAAACAGGATGTTGTCCAA -3'

ATGGAGAACTGTTCTGCTGCATCGACGTTCCTGACTGACAGCTTAGAGCTGGAGCTGGGG
ACGGAATGGTGCAAACCTCCTTACTTTTCTTGTGCTGTTGACAACAGAGGAGGAGGAAAA
CATTTTTCTGGAGAATCCTACCTCTGCAGCGGAGCCCTTAAGCGATTGATTTTGAATCTT
GACCCTTTACCAACTAATTTTGAAGAAGATACCTTGGAAATATTTGGCATTCAGTGGGTT
ACTGAAACAGCATTAGTGAATTCATCTAGAGAACTCTTTCATTTATTCAGGCAACAACTG

TACAACTTGGAAACCTTGTTACAGTCCAGTTGTGATTTTGGGAAGGTATCAACTCTACAC
TGCAAAGCAGACAATATTAGGCAGCAGTGTGTACTATTTCTCCATTATGTTAAAGTTTTC
ATCTTCAGGTATCTGAAAGTACAGAATGCTGAGAGTCATGTTCCTGTCCATCCTTATGAG
GCTTTGGAGGCTCAGCTTCCCTCAGTGTTGATTGATGAGCTTCATGGATTACTCTTGTAT
ATTGGACACCTATCTGAACTTCCCAGTGTTAATATAGGAGCATTTGTAAATCAAAACCAG
ATTAAGCTTTTTCCACCGTCATGGCATTTATTACATCTCCACTTGGATATACATTGGCTG
GTGCTAGAAATTCTTTACATGCTGGGTGAAAAATTGAAACAAGTTGTATATGGTCATCA
GTTTATGAATCTGGCAAGTGACAATTTAACCAACATCAGCCTATTTGAAGAACATTGTGA
AACTCTCCTTTGTGATTTAATAAGCCTGTCACTCAACAGGTACGACAAGGTTAGGTCTTC
TGAATCATTAATGAGTGACCAGTGTCCATGTTTATGCATTAAAGAATTATGGGTTCTACT
TATTCATCTTCTAGACCACAGAAGTAAATGGTTTGTCTCGGAATCATTTTGGAACTGGTT
GAATAAACTACTTAAAACACTGCTTGAAAAATCAAGTGACCGAAGAAGATCCTCTATGCC
TGTAATCCAGTCCAGGGATCCATTAGGTTTTAGTTGGTGGATTATTACTCATGTAGCATC
ATTTTACAAGTTTGATCGCCATGGAGTACCAGATGAAATGAGAAAAGTGGAATCAAATT
GGAACTTTGTAGAAGAACTGCTGAAAAAGTCCATCAGTGTTCAGGGTGTCATTCTAGAAG
AACAATTACGAATGTATCTTCACTGTTGTTTGACACTTTGTGATTTCTGGGAGCCAAACA
TTGCAATTGTTACCATTTTATGGGAATATTATAGTAAGAACCTGAATAGTTCCTTCAGTA
TTTCTTGGCTTCCTTTTAAAGGCCTTGCTAATACCATGAAGTCACCCTTGTCTATGCTTG
AAATGGTGAAGACTTGCTGTTGCGATAAACAAGATCAGGAACTATATAAATCCAGCAGT
AGTTATACTATTTTTCTTTGTATTCTGGCAAAAGTTGTTAAAAAAGCAATGAAGAGCAA
TGGCCCTCATCCTTGGAAACAAGTCAAAGGAAGAATATATTCAAAATTCCATCAAAAAAG
AATGGAAGAACTAACTGAAGTTGGTCTACAGAACTTTTTTAGCCTTTTTCTACTGTTAGC
AGCTGTTGCAGAGGTAGAAGATGTTGCAAGTCATGTTTTAGACCTCCTGAATTTCCTCAA
GCCTGCTTTTGTAATGTCTCAGAGAGCCCTCATTTGGAAGGGTCACATGGCCTTCCTCTT
GATGTATGCCCAGAAAAATCTGGACATTGGTGTTTTGGCTGAGAAATTTTCATGTGCTTT
CCGGGAGAAAGCAAAGGAATTCTTGGTGTCTAAGAATGAGGAAATGGTACAGAGACAGA
CTATCTGGACCCTTCTTTCCATATACATTGATGGTGTTCAAGAAGTGTTTGAGACCAGCT
ATTGCTTGTATCCTTCCCATGAAAAACTGCTTAATGATGGATTTAGTATGCTTCTGCGAG
CATGTCGAGAATCTGAACTTAGGACAGTATTGAGCTTCCTACAAGCTGTTCTGGCCAGAA
TCAGGAGTATGCATCAACAATTGTGTCAGGAACTTCAAAGGGACAATGTGGACCTATTTG
TACAGTCTTCATTATCGGCTAAAGAGCGCCACCTTGCTGCAGTTGCCAGTGCACTGTGGA
GACATTTCTTTTCATTTTTGAAGAGTCAGAGAATGTCACAGGTAGTGCCTTTCTCACAAC
TTGCGGATGCAGCTGCAGACTTTACTTTGCTAGCAATGGACATGCCAAGCACAGCTCCAT
CAGATTTTCAGCCTCAGCCAGTTATATCAATTATTCAACTTTTTGGTTGGGATGATATCA
TCTGCCCTCAAGTTGTAGCAAGATATTTAAGTCATGTCCTACAAAATAGCACATTATGTG
AAGCACTTTCTCATTCAGGCTATGTATCTTTTCAAGCCTTAACCGTAAGATCATGGATTC
GTTGTGTTTTGCAAATGTATATTAAAAACCTCTCTGGGCCTGATGATTTGCTCATAGATA
AAAATCTGGAAGAGGCAGTTGAAAAAGAGTACATGAAACAGTTGGTCAAACTGACAAGA
TTACTATTTAATCTCTCAGAAGTAAAGAGTATTTTCTCAAAGGCCCAAGCTGAATATTTA
TCCATCTCAGAAGACCCTAAAAAAGCACTTGTTCGATTCTTTGAGGCTGTTGGTGTAACT
TACGGGAACGTCCAGACACTTTCTGATAAATCTGCCATGGTCACAAAGTCCTTGGAATAC
CTTGGTGAAGTATTAAAATATATTAAGCCTTATTTGGGAAAAAAAGTTTTCAGTGCAGG
GCTGCAGCTGACTTATGGAATGATGGGAATTCTTGTGAAATCATGGGCACAAATCTTTGC
CACTTCTAAAGCCCAAAAATTACTATTCCGGATCATAGATTGTTTACTGCTGCCACATGC
AGTATTACAGCAAGAGAAGGAACTGCCTGCACCTATGTTGTCAGCAATTCAGAAAAGTCT
TCCTTTGTATCTCCAGGGCATGTGTATCGTGTGTTGTCAATCTCAAAATCCGAATGCCTA
TTTGAATCAATTGCTAGGGAATGTTATTGAGCAGTATATTGGGCGATTTCTTCCAGCTTC
ACCATATGTTTCAGATCTTGGACAACATCCTGTTTTG

9) **C-300 (C3)**



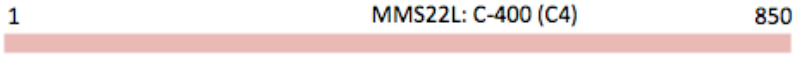
SF-P1: 5'-CTACTA GGGCCC ATGGAGAACTGTTCTGCTGC -3'

SR-C3: 5'-CTTCTT CTCGAG TTACTTATCGTCGTCATCCTTGTAATC
CACAAGAATTCCCATCATTC-3'

ATGGAGAACTGTTCTGCTGCATCGACGTTCCTGACTGACAGCTTAGAGCTGGAGCTGGGG
ACGGAATGGTGCAAACCTCCTTACTTTTCTTGTGCTGTTGACAACAGAGGAGGAGGAAAA
CATTTTTCTGGAGAATCCTACCTCTGCAGCGGAGCCCTTAAGCGATTGATTTTGAATCTT
GACCCTTTACCAACTAATTTTGAAGAAGATACCTTGGAAATATTTGGCATTCAGTGGGTT
ACTGAAACAGCATTAGTGAATTCATCTAGAGAACTCTTTCATTTATTCAGGCAACAACTG
TACAACTTGGAAACCTTGTTACAGTCCAGTTGTGATTTTGGGAAGGTATCAACTCTACAC
TGCAAAGCAGACAATATTAGGCAGCAGTGTGTACTATTTCTCCATTATGTTAAAGTTTTC
ATCTTCAGGTATCTGAAAGTACAGAATGCTGAGAGTCATGTTCCTGTCCATCCTTATGAG
GCTTTGGAGGCTCAGCTTCCCTCAGTGTTGATTGATGAGCTTCATGGATTACTCTTGTAT
ATTGGACACCTATCTGAACTTCCCAGTGTTAATATAGGAGCATTTGTAAATCAAAACCAG
ATTAAGCTTTTTCCACCGTCATGGCATTTATTACATCTCCACTTGGATATACATTGGCTG
GTGCTAGAAATTCTTTACATGCTGGGTGAAAAATTGAAACAAGTTGTATATGGTCATCA
GTTTATGAATCTGGCAAGTGACAATTTAACCAACATCAGCCTATTTGAAGAACATTGTGA
AACTCTCCTTTGTGATTTAATAAGCCTGTCACTCAACAGGTACGACAAGGTTAGGTCTTC
TGAATCATTAATGAGTGACCAGTGTCCATGTTTATGCATTAAAGAATTATGGGTTCTACT
TATTCATCTTCTAGACCACAGAAGTAAATGGTTTGTCTCGGAATCATTTTGGAACTGGTT
GAATAAACTACTTAAAACACTGCTTGAAAAATCAAGTGACCGAAGAAGATCCTCTATGCC
TGTAATCCAGTCCAGGGATCCATTAGGTTTTAGTTGGTGGATTATTACTCATGTAGCATC
ATTTTACAAGTTTGATCGCCATGGAGTACCAGATGAAATGAGAAAAGTGGAATCAAATT
GGAACTTTGTAGAAGAACTGCTGAAAAAGTCCATCAGTGTTCAGGGTGTCATTCTAGAAG
AACAATTACGAATGTATCTTCACTGTTGTTTGACACTTTGTGATTTCTGGGAGCCAAACA
TTGCAATTGTTACCATTTTATGGGAATATTATAGTAAGAACCTGAATAGTTCCTTCAGTA
TTTCTTGGCTTCCTTTTAAAGGCCTTGCTAATACCATGAAGTCACCCTTGTCTATGCTTG
AAATGGTGAAGACTTGCTGTTGCGATAAACAAGATCAGGAACTATATAAATCCAGCAGT
AGTTATACTATTTTTCTTTGTATTCTGGCAAAAGTTGTTAAAAAAGCAATGAAGAGCAA
TGGCCCTCATCCTTGGAAACAAGTCAAAGGAAGAATATATTCAAAATTCCATCAAAAAAG
AATGGAAGAACTAACTGAAGTTGGTCTACAGAACTTTTTTAGCCTTTTTCTACTGTTAGC
AGCTGTTGCAGAGGTAGAAGATGTTGCAAGTCATGTTTTAGACCTCCTGAATTTCCTCAA
GCCTGCTTTTGTAATGTCTCAGAGAGCCCTCATTTGGAAGGGTCACATGGCCTTCCTCTT
GATGTATGCCCAGAAAAATCTGGACATTGGTGTTTTGGCTGAGAAATTTTCATGTGCTTT
CCGGGAGAAAGCAAAGGAATTCTTGGTGTCTAAGAATGAGGAAATGGTACAGAGACAGA
CTATCTGGACCCTTCTTTCCATATACATTGATGGTGTTCAAGAAGTGTTTGAGACCAGCT
ATTGCTTGTATCCTTCCCATGAAAAACTGCTTAATGATGGATTTAGTATGCTTCTGCGAG
CATGTCGAGAATCTGAACTTAGGACAGTATTGAGCTTCCTACAAGCTGTTCTGGCCAGAA
TCAGGAGTATGCATCAACAATTGTGTCAGGAACTTCAAAGGGACAATGTGGACCTATTTG
TACAGTCTTCATTATCGGCTAAAGAGCGCCACCTTGCTGCAGTTGCCAGTGCACTGTGGA
GACATTTCTTTTCATTTTTGAAGAGTCAGAGAATGTCACAGGTAGTGCCTTTCTCACAAC
TTGCGGATGCAGCTGCAGACTTTACTTTGCTAGCAATGGACATGCCAAGCACAGCTCCAT

CAGATTTTCAGCCTCAGCCAGTTATATCAATTATTCAACTTTTTGGTTGGGATGATATCA
TCTGCCCTCAAGTTGTAGCAAGATATTTAAGTCATGTCCTACAAAATAGCACATTATGTG
AAGCACTTTCTCATTCAGGCTATGTATCTTTTCAAGCCTTAACCGTAAGATCATGGATTC
GTTGTGTTTTGCAAATGTATATTAAAAACCTCTCTGGGCCTGATGATTTGCTCATAGATA
AAAATCTGGAAGAGGCAGTTGAAAAAGAGTACATGAAACAGTTGGTCAAACTGACAAGA
TTACTATTTAATCTCTCAGAAGTAAAGAGTATTTTCTCAAAGGCCCAAGCTGAATATTTA
TCCATCTCAGAAGACCCTAAAAAAGCACTTGTTCGATTCTTTGAGGCTGTTGGTGTAACT
TACGGGAACGTCCAGACACTTTCTGATAAATCTGCCATGGTCACAAAGTCCTTGGAATAC
CTTGGTGAAGTATTAAAATATATTAAGCCTTATTTGGGAAAAAAAGTTTTCAGTGCAGG
GCTGCAGCTGACTTATGGAATGATGGGAATTCTTGTG

10) **C-400 (C4)**



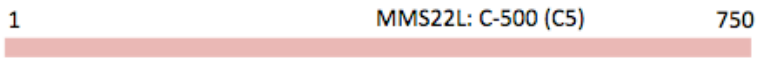
SF-P1: 5'-CTACTA GGGCCC ATGGAGAACTGTTCTGCTGC -3'

SR-C4: 5'-CTTCTT CTCGAG TTACTTATCGTCGTCATCCTTGTAATC
CATGTACTCTTTTTCAACTG -3'

ATGGAGAACTGTTCTGCTGCATCGACGTTCCTGACTGACAGCTTAGAGCTGGAGCTGGGG
ACGGAATGGTGCAAACCTCCTTACTTTTCTTGTGCTGTTGACAACAGAGGAGGAGGAAAA
CATTTTTCTGGAGAATCCTACCTCTGCAGCGGAGCCCTTAAGCGATTGATTTTGAATCTT
GACCCTTTACCAACTAATTTTGAAGAAGATACCTTGGAAATATTTGGCATTCAGTGGGTT
ACTGAAACAGCATTAGTGAATTCATCTAGAGAACTCTTTCATTTATTCAGGCAACAACTG
TACAACTTGGAAACCTTGTTACAGTCCAGTTGTGATTTTGGGAAGGTATCAACTCTACAC
TGCAAAGCAGACAATATTAGGCAGCAGTGTGTACTATTTCTCCATTATGTTAAAGTTTTC
ATCTTCAGGTATCTGAAAGTACAGAATGCTGAGAGTCATGTTCCTGTCCATCCTTATGAG
GCTTTGGAGGCTCAGCTTCCCTCAGTGTTGATTGATGAGCTTCATGGATTACTCTTGTAT
ATTGGACACCTATCTGAACTTCCCAGTGTTAATATAGGAGCATTTGTAAATCAAAACCAG
ATTAAGCTTTTTCCACCGTCATGGCATTTATTACATCTCCACTTGGATATACATTGGCTG
GTGCTAGAAATTCTTTACATGCTGGGTGAAAAATTGAAACAAGTTGTATATGGTCATCA
GTTTATGAATCTGGCAAGTGACAATTTAACCAACATCAGCCTATTTGAAGAACATTGTGA
AACTCTCCTTTGTGATTTAATAAGCCTGTCACTCAACAGGTACGACAAGGTTAGGTCTTC
TGAATCATTAATGAGTGACCAGTGTCCATGTTTATGCATTAAAGAATTATGGGTTCTACT
TATTCATCTTCTAGACCACAGAAGTAAATGGTTTGTCTCGGAATCATTTTGGAACTGGTT
GAATAAACTACTTAAAACACTGCTTGAAAAATCAAGTGACCGAAGAAGATCCTCTATGCC
TGTAATCCAGTCCAGGGATCCATTAGGTTTTAGTTGGTGGATTATTACTCATGTAGCATC
ATTTTACAAGTTTGATCGCCATGGAGTACCAGATGAAATGAGAAAAGTGGAATCAAATT
GGAACTTTGTAGAAGAACTGCTGAAAAAGTCCATCAGTGTTCAGGGTGTCATTCTAGAAG
AACAATTACGAATGTATCTTCACTGTTGTTTGACACTTTGTGATTTCTGGGAGCCAAACA
TTGCAATTGTTACCATTTTATGGGAATATTATAGTAAGAACCTGAATAGTTCCTTCAGTA
TTTCTTGGCTTCCTTTTAAAGGCCTTGCTAATACCATGAAGTCACCCTTGTCTATGCTTG
AAATGGTGAAGACTTGCTGTTGCGATAAACAAGATCAGGAACTATATAAATCCAGCAGT

AGTTATACTATTTTTCTTTGTATTCTGGCAAAAGTTGTTAAAAAAGCAATGAAGAGCAA
TGGCCCTCATCCTTGGAAACAAGTCAAAGGAAGAATATATTCAAAATTCCATCAAAAAAG
AATGGAAGAACTAACTGAAGTTGGTCTACAGAACTTTTTTAGCCTTTTTCTACTGTTAGC
AGCTGTTGCAGAGGTAGAAGATGTTGCAAGTCATGTTTTAGACCTCCTGAATTTCCTCAA
GCCTGCTTTTGTAATGTCTCAGAGAGCCCTCATTTGGAAGGGTCACATGGCCTTCCTCTT
GATGTATGCCCAGAAAAATCTGGACATTGGTGTTTTGGCTGAGAAATTTTCATGTGCTTT
CCGGGAGAAAGCAAAGGAATTCTTGGTGTCTAAGAATGAGGAAATGGTACAGAGACAGA
CTATCTGGACCCTTCTTTCCATATACATTGATGGTGTTCAAGAAGTGTTTGAGACCAGCT
ATTGCTTGTATCCTTCCCATGAAAAACTGCTTAATGATGGATTTAGTATGCTTCTGCGAG
CATGTCGAGAATCTGAACTTAGGACAGTATTGAGCTTCCTACAAGCTGTTCTGGCCAGAA
TCAGGAGTATGCATCAACAATTGTGTCAGGAACTTCAAAGGGACAATGTGGACCTATTTG
TACAGTCTTCATTATCGGCTAAAGAGCGCCACCTTGCTGCAGTTGCCAGTGCACTGTGGA
GACATTTCTTTTCATTTTTGAAGAGTCAGAGAATGTCACAGGTAGTGCCTTTCTCACAAC
TTGCGGATGCAGCTGCAGACTTTACTTTGCTAGCAATGGACATGCCAAGCACAGCTCCAT
CAGATTTTCAGCCTCAGCCAGTTATATCAATTATTCAACTTTTTGGTTGGGATGATATCA
TCTGCCCTCAAGTTGTAGCAAGATATTTAAGTCATGTCCTACAAAATAGCACATTATGTG
AAGCACTTTCTCATTCAGGCTATGTATCTTTTCAAGCCTTAACCGTAAGATCATGGATTC
GTTGTGTTTTGCAAATGTATATTAAAAACCTCTCTGGGCCTGATGATTTGCTCATAGATA
AAAATCTGGAAGAGGCAGTTGAAAAAGAGTACATG

11) **C-500 (C5)**



SF-P1: 5'-CTACTA GGGCCC ATGGAGAACTGTTCTGCTGC -3'

SR-P750: 5'-CTTCTT CTCGAG TTACTTATCGTCGTCATCCTTGTAATC
TGCTAGCAAAGTAAAGTCTG -3'

ATGGAGAACTGTTCTGCTGCATCGACGTTCCTGACTGACAGCTTAGAGCTGGAGCTGGGG
ACGGAATGGTGCAAACCTCCTTACTTTTCTTGTGCTGTTGACAACAGAGGAGGAGGAAAA
CATTTTTCTGGAGAATCCTACCTCTGCAGCGGAGCCCTTAAGCGATTGATTTTGAATCTT
GACCCTTTACCAACTAATTTTGAAGAAGATACCTTGGAAATATTTGGCATTCAGTGGGTT
ACTGAAACAGCATTAGTGAATTCATCTAGAGAACTCTTTCATTTATTCAGGCAACAACTG
TACAACTTGGAAACCTTGTTACAGTCCAGTTGTGATTTTGGGAAGGTATCAACTCTACAC
TGCAAAGCAGACAATATTAGGCAGCAGTGTGTACTATTTCTCCATTATGTTAAAGTTTTC
ATCTTCAGGTATCTGAAAGTACAGAATGCTGAGAGTCATGTTCCTGTCCATCCTTATGAG
GCTTTGGAGGCTCAGCTTCCCTCAGTGTTGATTGATGAGCTTCATGGATTACTCTTGTAT
ATTGGACACCTATCTGAACTTCCCAGTGTTAATATAGGAGCATTTGTAAATCAAAACCAG
ATTAAGCTTTTTCCACCGTCATGGCATTTATTACATCTCCACTTGGATATACATTGGCTG
GTGCTAGAAATTCTTTACATGCTGGGTGAAAAATTGAAACAAGTTGTATATGGTCATCA
GTTTATGAATCTGGCAAGTGACAATTTAACCAACATCAGCCTATTTGAAGAACATTGTGA
AACTCTCCTTTGTGATTTAATAAGCCTGTCACTCAACAGGTACGACAAGGTTAGGTCTTC
TGAATCATTAATGAGTGACCAGTGTCCATGTTTATGCATTAAAGAATTATGGGTTCTACT
TATTCATCTTCTAGACCACAGAAGTAAATGGTTTGTCTCGGAATCATTTTGGAACTGGTT
GAATAAACTACTTAAAACACTGCTTGAAAAATCAAGTGACCGAAGAAGATCCTCTATGCC

TGTAATCCAGTCCAGGGATCCATTAGGTTTTAGTTGGTGGATTATTACTCATGTAGCATC
ATTTTACAAGTTTGATCGCCATGGAGTACCAGATGAAATGAGAAAAGTGGAATCAAATT
GGAACTTTGTAGAAGAACTGCTGAAAAAGTCCATCAGTGTTCAGGGTGTCATTCTAGAAG
AACAATTACGAATGTATCTTCACTGTTGTTTGACACTTTGTGATTTCTGGGAGCCAAACA
TTGCAATTGTTACCATTTTATGGGAATATTATAGTAAGAACCTGAATAGTTCCTTCAGTA
TTTCTTGGCTTCCTTTTAAAGGCCTTGCTAATACCATGAAGTCACCCTTGTCTATGCTTG
AAATGGTGAAGACTTGCTGTTGCGATAAACAAGATCAGGAACTATATAAATCCAGCAGT
AGTTATACTATTTTTCTTTGTATTCTGGCAAAAGTTGTTAAAAAAGCAATGAAGAGCAA
TGGCCCTCATCCTTGGAAACAAGTCAAAGGAAGAATATATTCAAAATTCCATCAAAAAAG
AATGGAAGAACTAACTGAAGTTGGTCTACAGAACTTTTTTAGCCTTTTTCTACTGTTAGC
AGCTGTTGCAGAGGTAGAAGATGTTGCAAGTCATGTTTTAGACCTCCTGAATTTCCTCAA
GCCTGCTTTTGTAATGTCTCAGAGAGCCCTCATTTGGAAGGGTCACATGGCCTTCCTCTT
GATGTATGCCCAGAAAAATCTGGACATTGGTGTTTTGGCTGAGAAATTTTCATGTGCTTT
CCGGGAGAAAGCAAAGGAATTCTTGGTGTCTAAGAATGAGGAAATGGTACAGAGACAGA
CTATCTGGACCCTTCTTTCCATATACATTGATGGTGTTCAAGAAGTGTTTGAGACCAGCT
ATTGCTTGTATCCTTCCCATGAAAAACTGCTTAATGATGGATTTAGTATGCTTCTGCGAG
CATGTCGAGAATCTGAACTTAGGACAGTATTGAGCTTCCTACAAGCTGTTCTGGCCAGAA
TCAGGAGTATGCATCAACAATTGTGTCAGGAACTTCAAAGGGACAATGTGGACCTATTTG
TACAGTCTTCATTATCGGCTAAAGAGCGCCACCTTGCTGCAGTTGCCAGTGCACTGTGGA
GACATTTCTTTTCATTTTTGAAGAGTCAGAGAATGTCACAGGTAGTGCCTTTCTCACAAC
TTGCGGATGCAGCTGCAGACTTTACTTTGCTAGCA

12) **Full length MMS22L: F1034A (FL-FA only)**

SF-P1: 5'-CTACTA GGGCCC ATGGAGAACTGTTCTGCTGC -3'

SR-P1243: 5'-CTTCTT CTCGAG TTACTTATCGTCGTCATCCTTGTAATC
AGTATTATCATTTTCCAGTC -3'

Site mutagenesis primers:
Forward: **F-M22-FA**
5' - TTGAGCAGTATATTGGGCGAGCTCTTCCAGCTTCACCATATGT -3'

Reverse (complementary): **R-M22-FA**
5'- ACATATGGTGAAGCTGGAAGAGCTCGCCCAATATACTGCTCAA - 3'

ATGGAGAACTGTTCTGCTGCATCGACGTTCCTGACTGACAGCTTAGAGCTGGAGCTGGGG
ACGGAATGGTGCAAACCTCCTTACTTTTCTTGTGCTGTTGACAACAGAGGAGGAGGAAAA
CATTTTTCTGGAGAATCCTACCTCTGCAGCGGAGCCCTTAAGCGATTGATTTTGAATCTT
GACCCTTTACCAACTAATTTTGAAGAAGATACCTTGGAAATATTTGGCATTCAGTGGGTT
ACTGAAACAGCATTAGTGAATTCATCTAGAGAACTCTTTCATTTATTCAGGCAACAACTG
TACAACTTGGAAACCTTGTTACAGTCCAGTTGTGATTTTGGGAAGGTATCAACTCTACAC
TGCAAAGCAGACAATATTAGGCAGCAGTGTGTACTATTTCTCCATTATGTTAAAGTTTTC
ATCTTCAGGTATCTGAAAGTACAGAATGCTGAGAGTCATGTTCCTGTCCATCCTTATGAG
GCTTTGGAGGCTCAGCTTCCCTCAGTGTTGATTGATGAGCTTCATGGATTACTCTTGTAT

ATTGGACACCTATCTGAACTTCCCAGTGTTAATATAGGAGCATTTGTAAATCAAAACCAG
ATTAAGCTTTTTCCACCGTCATGGCATTTATTACATCTCCACTTGGATATACATTGGCTG
GTGCTAGAAATTCTTTACATGCTGGGTGAAAAATTGAAACAAGTTGTATATGGTCATCA
GTTTATGAATCTGGCAAGTGACAATTTAACCAACATCAGCCTATTTGAAGAACATTGTGA
AACTCTCCTTTGTGATTTAATAAGCCTGTCACTCAACAGGTACGACAAGGTTAGGTCTTC
TGAATCATTAATGAGTGACCAGTGTCCATGTTTATGCATTAAAGAATTATGGGTTCTACT
TATTCATCTTCTAGACCACAGAAGTAAATGGTTTGTCTCGGAATCATTTTGGAACTGGTT
GAATAAACTACTTAAAACACTGCTTGAAAAATCAAGTGACCGAAGAAGATCCTCTATGCC
TGTAATCCAGTCCAGGGATCCATTAGGTTTTAGTTGGTGGATTATTACTCATGTAGCATC
ATTTTACAAGTTTGATCGCCATGGAGTACCAGATGAAATGAGAAAAGTGGAATCAAATT
GGAACTTTGTAGAAGAACTGCTGAAAAAGTCCATCAGTGTTCAGGGTGTCATTCTAGAAG
AACAATTACGAATGTATCTTCACTGTTGTTTGACACTTTGTGATTTCTGGGAGCCAAACA
TTGCAATTGTTACCATTTTATGGGAATATTATAGTAAGAACCTGAATAGTTCCTTCAGTA
TTTCTTGGCTTCCTTTTAAAGGCCTTGCTAATACCATGAAGTCACCCTTGTCTATGCTTG
AAATGGTGAAGACTTGCTGTTGCGATAAACAAGATCAGGAACTATATAAATCCAGCAGT
AGTTATACTATTTTTCTTTGTATTCTGGCAAAAGTTGTTAAAAAAGCAATGAAGAGCAA
TGGCCCTCATCCTTGGAAACAAGTCAAAGGAAGAATATATTCAAAATTCCATCAAAAAAG
AATGGAAGAACTAACTGAAGTTGGTCTACAGAACTTTTTTAGCCTTTTTCTACTGTTAGC
AGCTGTTGCAGAGGTAGAAGATGTTGCAAGTCATGTTTTAGACCTCCTGAATTTCCTCAA
GCCTGCTTTTGTAATGTCTCAGAGAGCCCTCATTTGGAAGGGTCACATGGCCTTCCTCTT
GATGTATGCCCAGAAAAATCTGGACATTGGTGTTTTGGCTGAGAAATTTTCATGTGCTTT
CCGGGAGAAAGCAAAGGAATTCTTGGTGTCTAAGAATGAGGAAATGGTACAGAGACAGA
CTATCTGGACCCTTCTTTCCATATACATTGATGGTGTTCAAGAAGTGTTTGAGACCAGCT
ATTGCTTGTATCCTTCCCATGAAAAACTGCTTAATGATGGATTTAGTATGCTTCTGCGAG
CATGTCGAGAATCTGAACTTAGGACAGTATTGAGCTTCCTACAAGCTGTTCTGGCCAGAA
TCAGGAGTATGCATCAACAATTGTGTCAGGAACTTCAAAGGGACAATGTGGACCTATTTG
TACAGTCTTCATTATCGGCTAAAGAGCGCCACCTTGCTGCAGTTGCCAGTGCACTGTGGA
GACATTTCTTTTCATTTTTGAAGAGTCAGAGAATGTCACAGGTAGTGCCTTTCTCACAAC
TTGCGGATGCAGCTGCAGACTTTACTTTGCTAGCAATGGACATGCCAAGCACAGCTCCAT
CAGATTTTCAGCCTCAGCCAGTTATATCAATTATTCAACTTTTTGGTTGGGATGATATCA
TCTGCCCTCAAGTTGTAGCAAGATATTTAAGTCATGTCCTACAAAATAGCACATTATGTG
AAGCACTTTCTCATTCAGGCTATGTATCTTTTCAAGCCTTAACCGTAAGATCATGGATTC
GTTGTGTTTTGCAAATGTATATTAAAAACCTCTCTGGGCCTGATGATTTGCTCATAGATA
AAAATCTGGAAGAGGCAGTTGAAAAAGAGTACATGAAACAGTTGGTCAAACTGACAAGA
TTACTATTTAATCTCTCAGAAGTAAAGAGTATTTTCTCAAAGGCCCAAGCTGAATATTTA
TCCATCTCAGAAGACCCTAAAAAAGCACTTGTTCGATTCTTTGAGGCTGTTGGTGTAACT
TACGGGAACGTCCAGACACTTTCTGATAAATCTGCCATGGTCACAAAGTCCTTGGAATAC
CTTGGTGAAGTATTAAAATATATTAAGCCTTATTTGGGAAAAAAAGTTTTCAGTGCAGG
GCTGCAGCTGACTTATGGAATGATGGGAATTCTTGTGAAATCATGGGCACAAATCTTTGC
CACTTCTAAAGCCCAAAAATTACTATTCCGGATCATAGATTGTTTACTGCTGCCACATGC
AGTATTACAGCAAGAGAAGGAACTGCCTGCACCTATGTTGTCAGCAATTCAGAAAAGTCT
TCCTTTGTATCTCCAGGGCATGTGTATCGTGTGTTGTCAATCTCAAAATCCGAATGCCTA
TTTGAATCAATTGCTAGGGAATGTTATTGAGCAGTATATTGGGCGAGCTCTTCCAGCTTC
ACCATATGTTTCAGATCTTGGACAACATCCTGTTTTGCTGGCATTGAGAAACACAGCCAC
TATTCCGCCAATATCATCTCTAAAGAAATGCATTGTGCAAGTCATAAGGAAATCCTACCT
TGAGTATAAGGGGTCCTCACCTCCTCCTCGCTTAGCATCCATTCTGGCCTTCATCCTCCAA
CTCTTCAAGGAAACTAACACAGACATTTATGAAGTTGAACTACTCCTCCCTGGCATTTTA
AAATGCTTGGTGTTAGTCAGTGAACCACAAGTTAAAAGGCTGGCCACAGAGAACCTGCAA
TACATGGTAAAAGCCTGCCAAGTGGGGTCAGAAGAAGAACCTTCCTCCCAGCTGACTTCT

GTGTTTAGGCAGTTTATCCAGGATTATGGTATGAGGTACTATTACCAGGTTTACAGCATT
TTAGAAACAGTAGCAACATTGGACCAGCAGGTTGTCATCCACTTGATTTCTACCCTTACT
CAGTCTCTGAAGGATTCAGAGCAGAAATGGGGCCTTGGCAGGAATATAGCACAAAGGGA
AGCCTATAGCAAACTTTTGTCTCACCTTGGACAGATGGGACAAGATGAGATGCAGAGACT
GGAAAATGATAATACT

13) **S1** (**1-600aa**)

SF-P1: 5'-CTACTA GGGCCC ATGGAGAACTGTTCTGCTGC
SR-P600: 5'-CTTCTT
CTCGAG TTACTTATCGTCGTCATCCTTGTAATCCCGGAAAGCACATGAAAATT

ATGGAGAACTGTTCTGCTGCATCGACGTTCCTGACTGACAGCTTAGAGCTGGAGCTGGGGACGGA
ATGGTGCAAACCTCCTTACTTTTCTTGTGCTGTTGACAACAGAGGAGGAGGAAAACATTTTTCTG
GAGAATCCTACCTCTGCAGCGGAGCCCTTAAGCGATTGATTTTGAATCTTGACCCTTTACCAACTA
ATTTTGAAGAAGATACCTTGGAAATATTTGGCATTCAGTGGGTTACTGAAACAGCATTAGTGAAT
TCATCTAGAGAACTCTTTCATTTATTCAGGCAACAACTGTACAACTTGGAAACCTTGTTACAGTCC
AGTTGTGATTTTGGGAAGGTATCAACTCTACACTGCAAAGCAGACAATATTAGGCAGCAGTGTGT
ACTATTTCTCCATTATGTTAAAGTTTTCATCTTCAGGTATCTGAAAGTACAGAATGCTGAGAGTC
ATGTTCCTGTCCATCCTTATGAGGCTTTGGAGGCTCAGCTTCCCTCAGTGTTGATTGATGAGCTTC
ATGGATTACTCTTGTATATTGGACACCTATCTGAACTTCCCAGTGTTAATATAGGAGCATTTGTA
AATCAAAACCAGATTAAGCTTTTTCCACCGTCATGGCATTTATTACATCTCCACTTGGATATACAT
TGGCTGGTGCTAGAAATTCTTTACATGCTGGGTGAAAAATTGAAACAAGTTGTATATGGTCATCA
GTTTATGAATCTGGCAAGTGACAATTTAACCAACATCAGCCTATTTGAAGAACATTGTGAAACTC
TCCTTTGTGATTTAATAAGCCTGTCACTCAACAGGTACGACAAGGTTAGGTCTTCTGAATCATTA
ATGAGTGACCAGTGTCCATGTTTATGCATTAAAGAATTATGGGTTCTACTTATTCATCTTCTAGA
CCACAGAAGTAAATGGTTTGTCTCGGAATCATTTTGGAACTGGTTGAATAAACTACTTAAAACAC
TGCTTGAAAAATCAAGTGACCGAAGAAGATCCTCTATGCCTGTAATCCAGTCCAGGGATCCATTA
GGTTTTAGTTGGTGGATTATTACTCATGTAGCATCATTTTACAAGTTTGATCGCCATGGAGTACC
AGATGAAATGAGAAAAGTGGAATCAAATTGGAACTTTGTAGAAGAACTGCTGAAAAAGTCCATCA
GTGTTCAGGGTGTCATTCTAGAAGAACAATTACGAATGTATCTTCACTGTTGTTTGACACTTTGT
GATTTCTGGGAGCCAAACATTGCAATTGTTACCATTTTATGGGAATATTATAGTAAGAACCTGAA
TAGTTCCTTCAGTATTTCTTGGCTTCCTTTTAAAGGCCTTGCTAATACCATGAAGTCACCCTTGTC
TATGCTTGAAATGGTGAAGACTTGCTGTTGCGATAAACAAGATCAGGAACTATATAAATCCAGCA
GTAGTTATACTATTTTTCTTTGTATTCTGGCAAAAGTTGTTAAAAAAGCAATGAAGAGCAATGGC
CCTCATCCTTGGAAACAAGTCAAAGGAAGAATATATTCAAAATTCCATCAAAAAAGAATGGAAGA
ACTAACTGAAGTTGGTCTACAGAACTTTTTTAGCCTTTTTCTACTGTTAGCAGCTGTTGCAGAGGT
AGAAGATGTTGCAAGTCATGTTTTAGACCTCCTGAATTTCCTCAAGCCTGCTTTTGTAACGTCTCA
GAGAGCCCTCATTTGGAAGGGTCACATGGCCTTCCTCTTGATGTATGCCCAGAAAAATCTGGACAT
TGGTGTTTTGGCTGAGAAATTTTCATGTGCTTTCCGG

14) **S2 (601-1243aa)**

SF-P601: 5'-CTACTA GGGCCC GAGAAAGCAAAGGAATTCTT
SR-P1243: 5'-CTTCTT CTCGAG TTACTTATCGTCGTCATCCTTGTAATC
AGTATTATCATTTTCCAGTC -3

GAGAAAGCAAAGGAATTCTTGGTGTCTAAGAATGAGGAAATGGTACAGAGACAGACTATCTGGAC
CCTTCTTTCCATATACATTGATGGTGTTCAAGAAGTGTTTGAGACCAGCTATTGCTTGTATCCTTC
CCATGAAAAACTGCTTAATGATGGATTTAGTATGCTTCTGCGAGCATGTCGAGAATCTGAACTTA
GGACAGTATTGAGCTTCCTACAAGCTGTTCTGGCCAGAATCAGGAGTATGCATCAACAATTGTGT
CAGGAACTTCAAAGGGACAATGTGGACCTATTTGTACAGTCTTCATTATCGGCTAAAGAGCGCCA
CCTTGCTGCAGTTGCCAGTGCACTGTGGAGACATTTCTTTTCATTTTTGAAGAGTCAGAGAATGTC
ACAGGTAGTGCCTTTCTCACAACTTGCGGATGCAGCTGCAGACTTTACTTTGCTAGCAATGGACAT
GCCAAGCACAGCTCCATCAGATTTTCAGCCTCAGCCAGTTATATCAATTATTCAACTTTTTGGTTG
GGATGATATCATCTGCCCTCAAGTTGTAGCAAGATATTTAAGTCATGTCCTACAAAATAGCACAT
TATGTGAAGCACTTTCTCATTCAGGCTATGTATCTTTTCAAGCCTTAACCGTAAGATCATGGATTC
GTTGTGTTTTGCAAATGTATATTAAAAACCTCTCTGGGCCTGATGATTTGCTCATAGATAAAAAT
CTGGAAGAGGCAGTTGAAAAAGAGTACATGAAACAGTTGGTCAAACTGACAAGATTACTATTTAA
TCTCTCAGAAGTAAAGAGTATTTTCTCAAAGGCCCAAGCTGAATATTTATCCATCTCAGAAGACCC
TAAAAAAGCACTTGTTCGATTCTTTGAGGCTGTTGGTGTAACTTACGGGAACGTCCAGACACTTT
CTGATAAATCTGCCATGGTCACAAAGTCCTTGGAATACCTTGGTGAAGTATTAAAATATATTAAG
CCTTATTTGGGAAAAAAAGTTTTCAGTGCAGGGCTGCAGCTGACTTATGGAATGATGGGAATTCT
TGTGAAATCATGGGCACAAATCTTTGCCACTTCTAAAGCCCAAAAATTACTATTCCGGATCATAG
ATTGTTTACTGCTGCCACATGCAGTATTACAGCAAGAGAAGGAACTGCCTGCACCTATGTTGTCAG
CAATTCAGAAAAGTCTTCCTTTGTATCTCCAGGGCATGTGTATCGTGTGTTGTCAATCTCAAAATC
CGAATGCCTATTTGAATCAATTGCTAGGGAATGTTATTGAGCAGTATATTGGGCGATTTCTTCCA
GCTTCACCATATGTTTCAGATCTTGGACAACATCCTGTTTTGCTGGCATTGAGAAACACAGCCACT
ATTCCGCCAATATCATCTCTAAAGAAATGCATTGTGCAAGTCATAAGGAAATCCTACCTTGAGTA
TAAGGGGTCCTCACCTCCTCCTCGCTTAGCATCCATTCTGGCCTTCATCCTCCAACTCTTCAAGGA
AACTAACACAGACATTTATGAAGTTGAACTACTCCTCCCTGGCATTTTAAAATGCTTGGTGTTAG
TCAGTGAACCACAAGTTAAAAGGCTGGCCACAGAGAACCTGCAATACATGGTAAAAGCCTGCCAA
GTGGGGTCAGAAGAAGAACCTTCCTCCCAGCTGACTTCTGTGTTTAGGCAGTTTATCCAGGATTAT
GGTATGAGGTACTATTACCAGGTTTACAGCATTTTAGAAACAGTAGCAACATTGGACCAGCAGGT
TGTCATCCACTTGATTTCTACCCTTACTCAGTCTCTGAAGGATTCAGAGCAGAAATGGGGCCTTGG
CAGGAATATAGCACAAAGGGAAGCCTATAGCAAACTTTTGTCTCACCTTGGACAGATGGGACAAG
ATGAGATGCAGAGACTGGAAAATGATAATACT

15) **S3 (1-300aa)**

SF-P1: 5'-CTACTA GGGCCC ATGGAGAACTGTTCTGCTGC
SR-P300: 5'-CTTCTT CTCGAG TTACTTATCGTCGTCATCCTTGTAATC
AAGTAGAACCCATAATTCTT

ATGGAGAACTGTTCTGCTGCATCGACGTTCCTGACTGACAGCTTAGAGCTGGAGCTGGGGACGGA
ATGGTGCAAACCTCCTTACTTTTCTTGTGCTGTTGACAACAGAGGAGGAGGAAAACATTTTTCTG
GAGAATCCTACCTCTGCAGCGGAGCCCTTAAGCGATTGATTTTGAATCTTGACCCTTTACCAACTA
ATTTTGAAGAAGATACCTTGGAAATATTTGGCATTCAGTGGGTTACTGAAACAGCATTAGTGAAT
TCATCTAGAGAACTCTTTCATTTATTCAGGCAACAACTGTACAACTTGGAAACCTTGTTACAGTCC
AGTTGTGATTTTGGGAAGGTATCAACTCTACACTGCAAAGCAGACAATATTAGGCAGCAGTGTGT
ACTATTTCTCCATTATGTTAAAGTTTTCATCTTCAGGTATCTGAAAGTACAGAATGCTGAGAGTC

ATGTTCCTGTCCATCCTTATGAGGCTTTGGAGGCTCAGCTTCCCTCAGTGTTGATTGATGAGCTTC
ATGGATTACTCTTGTATATTGGACACCTATCTGAACTTCCCAGTGTTAATATAGGAGCATTTGTA
AATCAAAACCAGATTAAGCTTTTTCCACCGTCATGGCATTTATTACATCTCCACTTGGATATACAT
TGGCTGGTGCTAGAAATTCTTTACATGCTGGGTGAAAAATTGAAACAAGTTGTATATGGTCATCA
GTTTATGAATCTGGCAAGTGACAATTTAACCAACATCAGCCTATTTGAAGAACATTGTGAAACTC
TCCTTTGTGATTTAATAAGCCTGTCACTCAACAGGTACGACAAGGTTAGGTCTTCTGAATCATTA
ATGAGTGACCAGTGTCCATGTTTATGCATTAAAGAATTATGGGTTCTACTT

16) **S5 (600-900aa)**

SF-P600: 5'-CTACTA GGGCCC CGGGAGAAAGCAAAGGAATT
SR-P900: 5'-CTTCTT CTCGAG TTACTTATCGTCGTCATCCTTGTAATC
CCCGTAAGTTACACCAACAG

CGGGAGAAAGCAAAGGAATTCTTGGTGTCTAAGAATGAGGAAATGGTACAGAGACAGACTATCTG
GACCCTTCTTTCCATATACATTGATGGTGTTCAAGAAGTGTTTGAGACCAGCTATTGCTTGTATCC
TTCCCATGAAAAACTGCTTAATGATGGATTTAGTATGCTTCTGCGAGCATGTCGAGAATCTGAAC
TTAGGACAGTATTGAGCTTCCTACAAGCTGTTCTGGCCAGAATCAGGAGTATGCATCAACAATTG
TGTCAGGAACTTCAAAGGGACAATGTGGACCTATTTGTACAGTCTTCATTATCGGCTAAAGAGCG
CCACCTTGCTGCAGTTGCCAGTGCACTGTGGAGACATTTCTTTTCATTTTTGAAGAGTCAGAGAAT
GTCACAGGTAGTGCCTTTCTCACAACTTGCGGATGCAGCTGCAGACTTTACTTTGCTAGCAATGGA
CATGCCAAGCACAGCTCCATCAGATTTTCAGCCTCAGCCAGTTATATCAATTATTCAACTTTTTGG
TTGGGATGATATCATCTGCCCTCAAGTTGTAGCAAGATATTTAAGTCATGTCCTACAAAATAGCA
CATTATGTGAAGCACTTTCTCATTCAGGCTATGTATCTTTTCAAGCCTTAACCGTAAGATCATGGA
TTCGTTGTGTTTTGCAAATGTATATTAAAAACCTCTCTGGGCCTGATGATTTGCTCATAGATAAA
AATCTGGAAGAGGCAGTTGAAAAAGAGTACATGAAACAGTTGGTCAAACTGACAAGATTACTATT
TAATCTCTCAGAAGTAAAGAGTATTTTCTCAAAGGCCCAAGCTGAATATTTATCCATCTCAGAAG
ACCCTAAAAAAGCACTTGTTCGATTCTTTGAGGCTGTTGGTGTAACTTACGGG

17) **S6 (900-1243aa)**

SF-P900: 5'-CTACTA GGGCCC GGGAACGTCCAGACACTTTC
SR-P1243: 5'-CTTCTT CTCGAG TTACTTATCGTCGTCATCCTTGTAATC
AGTATTATCATTTTCCAGTC -3'

GGGAACGTCCAGACACTTTCTGATAAATCTGCCATGGTCACAAAGTCCTTGGAATACCTTGGTGA
AGTATTAAAATATATTAAGCCTTATTTGGGAAAAAAAGTTTTCAGTGCAGGGCTGCAGCTGACTT
ATGGAATGATGGGAATTCTTGTGAAATCATGGGCACAAATCTTTGCCACTTCTAAAGCCCAAAAA
TTACTATTCCGGATCATAGATTGTTTACTGCTGCCACATGCAGTATTACAGCAAGAGAAGGAACT
GCCTGCACCTATGTTGTCAGCAATTCAGAAAAGTCTTCCTTTGTATCTCCAGGGCATGTGTATCGT
GTGTTGTCAATCTCAAAATCCGAATGCCTATTTGAATCAATTGCTAGGGAATGTTATTGAGCAGT
ATATTGGGCGATTTCTTCCAGCTTCACCATATGTTTCAGATCTTGGACAACATCCTGTTTTGCTGG
CATTGAGAAACACAGCCACTATTCCGCCAATATCATCTCTAAAGAAATGCATTGTGCAAGTCATA
AGGAAATCCTACCTTGAGTATAAGGGGTCCTCACCTCCTCCTCGCTTAGCATCCATTCTGGCCTTC
ATCCTCCAACTCTTCAAGGAAACTAACACAGACATTTATGAAGTTGAACTACTCCTCCCTGGCATT
TTAAAATGCTTGGTGTTAGTCAGTGAACCACAAGTTAAAAGGCTGGCCACAGAGAACCTGCAATA
CATGGTAAAAGCCTGCCAAGTGGGGTCAGAAGAAGAACCTTCCTCCCAGCTGACTTCTGTGTTTAG
GCAGTTTATCCAGGATTATGGTATGAGGTACTATTACCAGGTTTACAGCATTTTAGAAACAGTAG

CAACATTGGACCAGCAGGTTGTCATCCACTTGATTTCTACCCTTACTCAGTCTCTGAAGGATTCAG
AGCAGAAATGGGGCCTTGGCAGGAATATAGCACAAAGGGAAGCCTATAGCAAACTTTTGTCTCAC
CTTGGACAGATGGGACAAGATGAGATGCAGAGACTGGAAAATGATAATACT

18) **S7 (150-450aa)**

SF-P150: 5'-CTACTA GGGCCC GCTGAGAGTCATGTTCCTGT
SR-P450: 5'-CTTCTT CTCGAG TTACTTATCGTCGTCATCCTTGTAATC
ATTAGCAAGGCCTTTAAAAG

GCTGAGAGTCATGTTCCTGTCCATCCTTATGAGGCTTTGGAGGCTCAGCTTCCCTCAGTGTTGATT
GATGAGCTTCATGGATTACTCTTGTATATTGGACACCTATCTGAACTTCCCAGTGTTAATATAGG
AGCATTTGTAAATCAAAACCAGATTAAGCTTTTTCCACCGTCATGGCATTTATTACATCTCCACTT
GGATATACATTGGCTGGTGCTAGAAATTCTTTACATGCTGGGTGAAAAATTGAAACAAGTTGTAT
ATGGTCATCAGTTTATGAATCTGGCAAGTGACAATTTAACCAACATCAGCCTATTTGAAGAACAT
TGTGAAACTCTCCTTTGTGATTTAATAAGCCTGTCACTCAACAGGTACGACAAGGTTAGGTCTTCT
GAATCATTAATGAGTGACCAGTGTCCATGTTTATGCATTAAAGAATTATGGGTTCTACTTATTCA
TCTTCTAGACCACAGAAGTAAATGGTTTGTCTCGGAATCATTTTGGAACTGGTTGAATAAACTAC
TTAAAACACTGCTTGAAAAATCAAGTGACCGAAGAAGATCCTCTATGCCTGTAATCCAGTCCAGG
GATCCATTAGGTTTTAGTTGGTGGATTATTACTCATGTAGCATCATTTTACAAGTTTGATCGCCA
TGGAGTACCAGATGAAATGAGAAAAGTGGAATCAAATTGGAACTTTGTAGAAGAACTGCTGAAA
AAGTCCATCAGTGTTCAGGGTGTCATTCTAGAAGAACAATTACGAATGTATCTTCACTGTTGTTT
GACACTTTGTGATTTCTGGGAGCCAAACATTGCAATTGTTACCATTTTATGGGAATATTATAGTA
AGAACCTGAATAGTTCCTTCAGTATTTCTTGGCTTCCTTTTAAAGGCCTTGCTAAT

19) **S8 (450-750aa)**

SF-P450: 5'-CTACTA GGGCCC AATACCATGAAGTCACCCTT
SR-P750: 5'-CTTCTT CTCGAG TTACTTATCGTCGTCATCCTTGTAATC
TGCTAGCAAAGTAAAGTCTG

AATACCATGAAGTCACCCTTGTCTATGCTTGAAATGGTGAAGACTTGCTGTTGCGATAAACAAGA
TCAGGAACTATATAAATCCAGCAGTAGTTATACTATTTTTCTTTGTATTCTGGCAAAAGTTGTTA
AAAAAGCAATGAAGAGCAATGGCCCTCATCCTTGGAAACAAGTCAAAGGAAGAATATATTCAAAA
TTCCATCAAAAAAGAATGGAAGAACTAACTGAAGTTGGTCTACAGAACTTTTTTAGCCTTTTTCT
ACTGTTAGCAGCTGTTGCAGAGGTAGAAGATGTTGCAAGTCATGTTTTAGACCTCCTGAATTTCC
TCAAGCCTGCTTTTGTAACGTCTCAGAGAGCCCTCATTTGGAAGGGTCACATGGCCTTCCTCTTGA
TGTATGCCCAGAAAAATCTGGACATTGGTGTTTTGGCTGAGAAATTTTCATGTGCTTTCCGGGAG
AAAGCAAAGGAATTCTTGGTGTCTAAGAATGAGGAAATGGTACAGAGACAGACTATCTGGACCCT
TCTTTCCATATACATTGATGGTGTTCAAGAAGTGTTTGAGACCAGCTATTGCTTGTATCCTTCCCA
TGAAAAACTGCTTAATGATGGATTTAGTATGCTTCTGCGAGCATGTCGAGAATCTGAACTTAGGA
CAGTATTGAGCTTCCTACAAGCTGTTCTGGCCAGAATCAGGAGTATGCATCAACAATTGTGTCAG
GAACTTCAAAGGGACAATGTGGACCTATTTGTACAGTCTTCATTATCGGCTAAAGAGCGCCACCTT
GCTGCAGTTGCCAGTGCACTGTGGAGACATTTCTTTTCATTTTTGAAGAGTCAGAGAATGTCACA
GGTAGTGCCTTTCTCACAACTTGCGGATGCAGCTGCAGACTTTACTTTGCTAGCA

20) **S9 (750-1050aa)**

SF-P750: 5'-CTACTA GGGCCC GCAATGGACATGCCAAGCAC
SR-P1050: 5'-CTTCTT CTCGAG TTACTTATCGTCGTCATCCTTGTAATC
CAAAACAGGATGTTGTCCAA

GCAATGGACATGCCAAGCACAGCTCCATCAGATTTTCAGCCTCAGCCAGTTATATCAATTATTCAA
CTTTTTGGTTGGGATGATATCATCTGCCCTCAAGTTGTAGCAAGATATTTAAGTCATGTCCTACA
AAATAGCACATTATGTGAAGCACTTTCTCATTCAGGCTATGTATCTTTTCAAGCCTTAACCGTAAG
ATCATGGATTCGTTGTGTTTTGCAAATGTATATTAAAAACCTCTCTGGGCCTGATGATTTGCTCA
TAGATAAAAATCTGGAAGAGGCAGTTGAAAAAGAGTACATGAAACAGTTGGTCAAACTGACAAG
ATTACTATTTAATCTCTCAGAAGTAAAGAGTATTTTCTCAAAGGCCCAAGCTGAATATTTATCCA
TCTCAGAAGACCCTAAAAAAGCACTTGTTCGATTCTTTGAGGCTGTTGGTGTAACTTACGGGAAC
GTCCAGACACTTTCTGATAAATCTGCCATGGTCACAAAGTCCTTGGAATACCTTGGTGAAGTATT
AAAATATATTAAGCCTTATTTGGGAAAAAAAGTTTTCAGTGCAGGGCTGCAGCTGACTTATGGAA
TGATGGGAATTCTTGTGAAATCATGGGCACAAATCTTTGCCACTTCTAAAGCCCAAAAATTACTA
TTCCGGATCATAGATTGTTTACTGCTGCCACATGCAGTATTACAGCAAGAGAAGGAACTGCCTGC
ACCTATGTTGTCAGCAATTCAGAAAAGTCTTCCTTTGTATCTCCAGGGCATGTGTATCGTGTGTTG
TCAATCTCAAAATCCGAATGCCTATTTGAATCAATTGCTAGGGAATGTTATTGAGCAGTATATTG
GGCGATTTCTTCCAGCTTCACCATATGTTTCAGATCTTGGACAACATCCTGTTTTG

6 ACKNOWLEDGEMENTS

First and foremost, I would like to thank Prof. Petr Cejka for giving me the opportunity to work on this project, for great supervision throughout the duration of my PhD and for helpful advice and ideas, which led to a successful outcome.

Also I would like to thank all members of Cejka lab for stimulating working atmosphere and valuable discussions. I would like to especially thank Lepakshi for considerable help at the beginning of my PhD project, which made my life way easier.

Special thanks goes to my committee members Prof. Massimo Lopes and Prof. Matthias Peter, Wojciech Piwko, and Pavel Janscak for their scientific advice and fruitful discussion during my PhD.

Many thanks to the IMCR for unforgettable time we have spent together either at work or outside. I would like to thank Farah and Odete for keeping the institute running smoothly and for their help. Many thanks to Josef Jiricny for leading such a great institute like IMCR.

I thank Cancer Biology PhD program for creating such an amazing opportunity to learn new things and interact with fellow PhD students.

And most importantly, I would like to thank my family for an amazing support throughout my studies, especially my mum, Vláďa, Kuba, and Honza for believing in me and pushing me always forward. Furthermore, I would like to thank my friends, especially Chris, Hannah, and Sebi for huge support in the past few months.

7 CURRICULUM VITAE

Name	**Lucie Jarmila Mlejnková**
Date of Birth	10.6.1987
Nationality	Czech

Education

01/2013 - present	**PhD in Cancer Biology**, Group of Prof. Petr Cejka, Institute of Molecular Cancer Research, University of Zurich, Switzerland. Title of PhD thesis: Elucidating the Role of MMS22L-TONSL Heterodimer in Homologous Recombination.
09/2011 - 10/2012	**MSc. with Distinction in Biotechnology**, Edinburgh Napier University, Scotland, UK. Title of Master thesis: *Caenorhabditis elegans* as a model in the assessment of ZnO and TiO_2 toxicity.
09/2008 - 08/2011	**Bc. in Biochemistry and Biotechnology**, Institute of Chemical Technology in Prague, Czech Republic Title of Bachelor thesis: Changes in *Caenorhabditis elegans* fat metabolism after inhibition of selected nuclear hormone receptors (nhr-68 and nhr-137).
06/2007	Gymnazium prof. Jana Patocky, Prague, Czech Republic Graduated with distinction.

Professional Experience

08/2007 - 08/2011	**Lab technician** in the Laboratory of Model Systems, Charles University in Prague, Czech Republic

Teaching Experience

04/2014 - 05/2014	Teaching assistant, practical block course, University of Zurich
03/2015 - 04/2015	Teaching assistant, practical block course, University of Zurich

10/2015 - 11/2015	Teaching assistant, Classical and Molecular Genetics, University of Zurich
03/2016 - 04/2016	Teaching assistant, practical block course, University of Zurich

Awards and Scholarship

10/2012	University Medal - awarded in recognition of outstanding performance throughout the programme of study.
08/2011 - 09/2012	Student Awards Agency for Scotland scholarship.
10/2010	Student's Scientific Conference for postgraduate students, Prague, Czech Republic: Abstract selected by Roche company.

Scientific Conferences

03/2016	**6th Student Cancer Retreat** in Davos, Switzerland: gave a talk entitled „Elucidating the role of MMS22L-TONSL heterodimer in homologous recombination“
09/2015	**Eukaryotic DNA Replication & Genome Maintenance** conference, Cold Spring Harbor Laboratory, USA: presented a poster (2015).
04/2015	**6th Cancer Network Zürich retreat** in Emmetten, Switzerland: presented a poster (2015).
01/2015	**Brupbacher Symposium** in Zürich, Switzerland: presented a poster (2015).
02/2014	**5th Student Cancer Retreat** in Filzbach, Switzerland: presented a poster (2014).
10/2010	**Student's Scientific Conference for postgraduate students,** Prague, Czech Republic.
07/2008	**Participation on poster:** Simeckova K., **Mlejnkova L.,** Pohludka M., Kostrouch Z. and Kostrouchova M. (2008) Functional analysis of *C. elegans* supplementary nuclear receptors characterized by the P box sequence CNGCKT. *European Worm Meeting.*

Publication

Giulia Saredi, Hongda Huang, Colin M. Hammond, Constance Alabert, Simon Bekker-Jensen, Ignasi Forne, Nazaret Reverón-Gómez, Benjamin M. Foster, **Lucie Mlejnkova**, Till Bartke, Petr Cejka, Niels Mailand, Axel Imhof, Dinshaw J. Patel, and Anja Groth (2016) H4K20me0 marks post-replicative chromatin and recruits the TONSL-MMS22L DNA repair complex, *Nature*.